



THE UNIVERSITY *of* EDINBURGH

Thesis scanned from best copy available: may contain faint or blurred text, and/or cropped or missing pages.

QUANTITY IN ICELANDIC:
A HISTORICAL AND COMPARATIVE STUDY

Kristján Árnason

A thesis submitted for the degree
of Doctor of Philosophy
University of Edinburgh

1977



ACKNOWLEDGMENTS.

A number of people deserve my thanks for help during the writing of this thesis, but surely none more than my supervisor, Dr. Roger Lass. Not only is he partly responsible for a good number of observations made in the following and, of course, the elimination of a probably no less substantial number of errors, but more importantly, he has given me exactly the type of encouragement I needed while working on this project. I think I can honestly say that I couldn't have wished for a better navigator on this perilous voyage.

A number of other people here in Edinburgh have been helpful to me: Gill Brown, John Anderson, Professor David Abercrombie and Hermann Pálsson have all read parts of earlier drafts and made suggestions for improvement. I am grateful to them.

I wish to express special thanks to my old teacher in Reykjavík, Hreinn Benediktsson, who has read earlier versions of some of the following chapters and whose influence on my work is perhaps only too obvious. I am also grateful to Stefán Karlsson of the Arnarnagmæan Institute in Reykjavík for helpful comments and to the Institute for allowing me to use its facilities on two occasions.

Mrs. Sheila Coppock has proof-read the final version of the thesis and eliminated a number of errors, and suggested improvements in style. I thank her warmly.

Grants from the British Council, the University of Edinburgh and the Icelandic Science Foundation (Vísindasjóður) combined to keep me free of financial worries during the course of my work here. I want to express my sincere gratitude to these institutions.

Finally, I would like to thank my wife and children for going into voluntary exile from their friends and country to stay with me these years in Edinburgh and my wife for typing the thesis with great skill on our thirty year old typewriter.

Department of Linguistics
University of Edinburgh

July 1977

Kristján Árnason

ABSTRACT

Chapter I is an introduction and statement of some basic phonological beliefs. A brief survey is given of the phonology of Modern Icelandic.

Chapter II. This is a study of the status and function of quantity in Modern Icelandic. It is suggested that vowel length is predictable on the basis of stress and the following consonantism. The phonological concept of syllable is discussed.

Chapter III is a survey of the function and history of quantity in the other Germanic languages.

Chapter IV. An investigation is made into the development of quantity in Icelandic from about 1200 down to modern times. A change took place from 'free' to predictable vowel length. The study involves investigation of the evidence given by the metrics of rímur- and dróttkvætt-poetry. It is concluded that the change from the old to the modern system was gradual. Concepts like 'cause of linguistic change' and 'explanation in historical linguistics' are discussed, and also the notion of 'historical conspiracy'.

Chapter V discusses ways of analysing length in phonological systems in general. It also touches on the relation between the subsystems of dichotomous systems split by quantity. The analyses of Old and Modern Icelandic are then reconsidered in the light of this discussion.

CONTENTS

Title page	1
Acknowledgements	ii
Abstract	iv
<u>Chapter I. Introduction</u>	1
1. Purpose	1
2. Phonological creed	4
3. A preview of Icelandic phonology	9
<u>Chapter II. Modern Icelandic</u>	17
1. The length rule	17
2. Preaspiration	31
3. Exceptions to the length rule, syllabification .	38
4. Stress	70
5. Some problematic forms	84
<u>Chapter III. Length in other Germanic languages</u> ...	89
1. Faroese	89
2. Norwegian	98
3. Swedish	110
4. Danish	118
5. Gothic	124
6. German	133
7. English	138
<u>Chapter IV. The development in Icelandic</u>	146
0. Introductory	146
0.1 The prosodic and segmental aspects of the problem	146
0.2 Sources of evidence	149
1. The phonology of Icelandic about 1200	150
1.1 The vowel system	150
1.2 The prosodic system	169
1.2.1 The evidence of metrics	169
1.2.2 Overlong syllables	180
1.3 Summary. The situation in the early 13th century	194

2.	From Old to Modern Icelandic	195
2.1	The term 'quantity shift'	197
2.2	Dating the changes	209
2.2.1	The evidence of metrics	211
2.2.1.1	<u>Dróttkvætt</u>	211
2.2.1.2	<u>Rímur</u>	221
2.2.1.3	The last feet of untruncated lines	236
2.3	The environment <u>p,t,k,s</u> + <u>v,i,r</u>	249
2.4	Conclusion	265
3.	Explaining the quantity shift	268
3.1	Explanation	269
3.2	The term 'cause' in historical linguistics	291
3.3	The causes of lengthening and shortening	298
3.4	A shift or a conspiracy?	306
	<u>Chapter V. The phonological analysis of length</u>	312
1.	Ways of accounting for length	312
2.	Types of correspondences between dichotomous systems	323
3.	Diphthongs and long vowels	335
4.	The phonological analysis of length in Old Icelandic	340
4.1	The nature of length about 1200	340
4.2	The status of length as an inherent feature	353
5.	The phonological analysis of length in Modern Icelandic, a quick second look	358
	<u>Footnotes</u>	364
	<u>Bibliography</u>	373

Chapter I. INTRODUCTION

1. Purpose

This study has, most broadly speaking, a twofold purpose: (a) It is intended to shed some light on the phonological status of length in Icelandic from about 1200 down to modern times, and (b) it is hoped that the problems that are discussed raise theoretical issues that are of interest in general linguistics, more particularly historical linguistics and phonology, as the state of the art is at present. As a side effect it is my hope that what follows (especially Chapter III) may create a useful framework for a more general study of the history of quantity and other prosodic features in the other Scandinavian and Germanic languages.

Chapter II is a study of how to deal with quantity in Modern Icelandic. It is proposed that length of vowels is predictable phonologically on the basis of stress and syllabification, in such a way that if more than one consonant follows the vowel within the same syllable the vowel is short, otherwise it is long. The syllabification that is proposed is connected with stress and is basically what may be called 'final maximalistic' in that as many consonants as possible are assigned to the 'code' of a stressed syllable. In particular it is suggested that some sort of phonotactic constraint prevails that forbids

syllables ending in a sequence of a /p/, /t/, /k/ or /s/ followed by a /v/, /j/ or /r/. This, it is suggested, accounts for the fact that forms like neþja [ne:þja] 'cold weather', have long vowels in spite of the fact that two consonants follow.

Chapter III gives short summaries of the histories of quantity in other Germanic languages. It is emphasized that the quantity development in Faroese, Norwegian and Swedish seems to be highly similar to what took place in Icelandic, whereas the developments in Danish and the West-Germanic languages are similar in many respects but different from the others. The most important difference between Icelandic, Faroese, Norwegian and Swedish on the one side and Danish, English and German on the other is that in the latter vowel length can (except for a few dialects, e.g. Scots and Upper Austrian) be said still to be distinctive in spite of the changes that took place and 'aimed at' making it predictable.

Chapter IV deals with the development in Icelandic in detail. It starts off in Section 1 by summarising what can be said about the situation in Icelandic about 1200 and then moves on in Section 2 to try to trace the changes that were to affect the Old Icelandic structure. Much of the chapter is devoted to the evaluation of metrics as evidence about the quantity structure of the language. Particular attention is paid to the evidence given by dróttkvætt- and rimur-poetry from the periods dealt with. The third and final section of this chapter is devoted to the problem of 'explaining' the changes that led to the reorganization of quantity that is usually termed the

Icelandic 'quantity shift' (Danish: kvantitetssvæltning-
en, Icelandic: hljóðvalarbreytingin). This change was
that length, having been 'free' in vowels in Old Iceland-
ic, came to be determined by the context in the way de-
scribed in Chapter II. It is suggested that stress was
crucially involved in the change (or changes) and that the
result was to produce a unit, central in the phonology of
Modern Icelandic, the 'stressed syllable'. Among the
theoretical questions touched on is the problem of what
can be called an explanation in historical linguistics
and what sort of metatheoretical demands should be made
on statements in general and in historical linguistics
in particular in order that they may be called scientific
statements or explanations. Closely connected with this,
the concept of 'cause of a linguistic change' is discuss-
ed. A third theoretical problem dealt with is the idea
of 'historical conspiracies' (cf. Lass 1974), and the
idea that changes may aim at a certain structure.

Chapter V, which is the final chapter of this study,
deals with the question of how, in general, quantity is
to be treated in phonological descriptions and how it can
function in phonological systems. Connected with this is
the question of how dichotomous vowel systems like those
split by quantity are organized, one subsystem in rela-
tion to the other. In the light of these general consi-
derations the structure of Old Icelandic is then reconsi-
dered, and in the final section a short commentary is made
on the situation in Modern Icelandic.

2. Phonological creed.

The following study is intended to be maximally free of phonological dogmas. It can perhaps be described as 'mildly generative'. The overall idea is that of some sort of generative phonology, with severe reservations about the power it has been given by its most orthodox practitioners. Although many of the issues concerning phonological theory, like the questions of abstractness, extrinsic rule ordering and distinctive features, are largely avoided, my general tendency is to believe that underlying representations should be as concrete as possible even though it may lead to complications in morphology (and syntax). In general I believe that many of the morphophonemic regularities that have been incorporated into the 'phonology' in works like Chomsky and Halle (1968) and ^{S.}Anderson (1974) don't belong in phonology at all, but should be dealt with in a separate morphophonemic component. My creed is probably similar in many ways to the ideas set forth in Linell (1974) and Derwing (1973), although the motivation for my creed may be slightly different from the force behind Linell's and Derwing's criticisms of generative phonology. Their main argument is the psychological implausibility of abstract morpheme invariants like /re=duke/ for the common core in reduce and reduction. Although I agree with this, it seems to me (as both Derwing and Linell admit) that a case can be made for some sort of common denominator for the [redʒuːs] of reduce and the [redʌk] of reduction. The thing is that the moti-

vation for it is by no means phonological, but something else, either semantic, morphological, or both. That is, somewhere in their grammar (their 'competence') most speakers probably have some 'device' that represents the fact that the forms reduce and reduction are related, but that device is, I think, not phonological. The most important reason for their being seen as having something in common is the fact that they are semantically and morphologically related.

Let's have a look at a few pairs in Icelandic which can be said to show a morphophonemic alternation between [ou] (orthographic ó) and [ai], (orthographic æ). A number of forms show inflectional alternations:

(a) fór - færi - farum 'went' (indicative sg. vs subjunctive sg. vs subjunctive pl.)

tók - tæki - tækjum 'took' - " -

stór - starri - starstur 'big' (positive vs comparative vs superlative)

Other forms seem to show derivational regularity:

(b) <u>óp</u> 'a cry'	<u>æpa</u> 'to cry'
<u>hól</u> 'praise, complement'	<u>hæla</u> 'to praise'
<u>blóð</u> 'blood'	<u>blæða</u> 'to bleed'

Lastly we have the following forms:

(c) <u>skór</u> 'a shoe'	<u>skæri</u> 'scissors'
<u>sól</u> 'sun'	<u>sæla</u> 'happiness'
<u>dós</u> 'a tin'	<u>dæsa</u> 'to sigh'

Let us now ask ourselves whether we should, in the spirit of 'orthodox' generative phonology (of the Chomsky-Halle (1968) type), set up underlying forms that both the

alternants with [ou] and [ai] can be derived from. First of all: why would we want to do this? The answer would be, in the case of the inflecting forms, that forms with the same 'lexical meaning' show regular alternations between [ou] and [ai] according to morphological environment. In the second set of examples, there is a similar alternation based on derivation, [ou] in the nouns, but [ai] in the verbs. I think it is fair to say that if this data were to have been handled by Chomsky and Halle (1968), they would have at least seriously considered the possibility of deriving the alternating forms in question from underlying invariant morphemes. But notice that from the point of view of this data, the motivation for this is not phonological at all, but semantic and morphological. The only place where phonological considerations (i.e. things to do with phonic substance and linguistic structure directly related to it) can conceivably be brought forth is in the forms, fari, taki and starri, where the [ai]-vowel is followed by a front vowel in the ending. [ai] is more 'frontish' than [ou] (if that can mean anything). But this does not work for farum, and starstar (it might for takium, since it has a palatal: [tʰai:cʰɪm] following the vowel), nor for spa, hala and blaða. If we now look at the last set of examples, we have the same sort of alternations in the stem vowels, but no semantic similarity between the forms. What is to be done here? I am sure that no generative 'phonologist' would suggest that these pairs be derived from a common underlying 'phonological'

form. Why? Because they are not semantically related. So, the less semantic motivation there is for setting up morpheme invariants, the less likely the generative 'phonologist' is to want to set up abstract 'systematic phonemic' forms.

I realize that in this battle against my generative phonologist 'straw-man' I am taking something for granted that many linguists would not subscribe to, namely that it is entirely clear what is phonological and what is not. Where does the border lie between phonology and morphology/morphophonemics? I can only say that it seems to me that there is a genuine difference between purely phonological processes like the palatalization of all velars preceding front vowels and a morphophonemic alternation like the one in fór - færum. The former can be related fairly directly to a phonetic process of coarticulation, whereas the alternation between [ou] and [ai] in the forms above can only be explained phonetically by going back, at least to Old Icelandic, and probably all the way back to Proto-Nordic. I don't have a definite answer (yet) to the question of what processes should be called genuinely phonological and which ones not, and where phonology ends and morphophonemics begins, but I am in no doubt that many of the things proposed by Chomsky and Halle and their followers as underlying morphemes and the rules allegedly connecting them with surface phonetic forms are such that calling them phonological and still maintaining that phonology should as such deal with the phonic structure of language is absurd.

As to the question of extrinsic rule ordering, I must admit that my beliefs are as ill-founded here as in the case of the abstractness problem, but I will state them anyway. I believe that phonological processes are not ordered by special ad hoc devices stating the relation between two (or more) specific rules, rather that whenever rules are 'ordered', they are so simply because the order in which they will apply is the only order in which they can apply, granted that a rule applies whenever its input-conditions are satisfied. In other words, I believe in intrinsic ordering, based on principles like Kiparsky's bleeding / feeding relation. I have a feeling that in arguments for the existence of some extrinsic ordering, be it 'local' (S. Anderson 1974) or 'global' (Cathey and Demers 1976 and Vago 1977) the, to my mind, venial sin of confusing phonology and morphophonemics is responsible.

As to the third major issue in phonology, the status, number and nature of distinctive features, I mention my beliefs on pp. 316-17 (Chapter V). I am very sceptical of the idea of universal phonological features of the type presented in Chomsky and Halle (1968) and Ladefoged (1971), and earlier in Jakobson, Fant and Halle (1952) and Jakobson and Halle (1956). I am particularly pessimistic as to their status as a part of an innate faculté de langage. I think distinctive features are language specific and that their apparent universality derives from the fact that all

human languages work with the same type of phonetic means, i.e. sound waves and human speech organs, and the limits on the number of distinctions that any particular language can make, as for example in vowel height, simply derive on the one hand from the fact that there is a limit to the different configurations of formants the human ear can in practice distinguish as different in every day speech, and on the other from the limits to the sounds that human articulatory organs can produce. As is the case with my other creeds in phonology outlined above, this one is basically a matter of faith, but there are arguments that can be put forth in support of it, although I won't attempt to present them here. (For some arguments see Sampson 1974 and p.317 below.)

3. A preview of Icelandic phonology

A considerable amount of work has been done on Modern Icelandic Phonology. General handbooks of Icelandic are e.g. Einarsson (1945) and Kelly and Kress (1972). General treatments of the phonology are to be found in Malone (1952), Haugen (1958) and the vowel phonology in particular is treated e.g. in Benediktsson (1959:301-302) and Steblin-Kamenskij (1960). Generative treatment of aspects of Icelandic phonology is to be found in S. Anderson (1969, 1972a and 1974). Works dealing specifically with quantity are: Malone (1953) Benediktsson (1963), Árnason (1975) Garnes (1973) and Kjartansson (1974). Among phonetic studies of Icelandic may be mentioned: Ófeigsson

(1920-24), Einarsson (1927), Bergsveinsson (1941), Guðfinnsson (1946 and 1964) and Pétursson (1974).

An important phonetico-phonological study of quantity in Modern Icelandic is to be found in Garnes (1974a), where results from various experiments are reported and commented on.

I will not write commentaries on any of these works; they will be referred to in what follows as occasioned by my discussion. Instead, I will try to give a reasonably clear and unbiased account of the most important features of Modern Icelandic phonology for the convenience of those who are not familiar with the data. Evidently some prejudices of mine will affect the following account, since I don't pay attention to all analyses proposed by all scholars, but I hope these prejudices are shared by the majority of people, in other words, I hope I am not presenting a minority view of Icelandic phonology.

Vowels:

Modern Icelandic has the following thirteen vowel phonemes, represented in a 'phonemicized' broad IPA transcription:

Monophthongs.		Diphthongs.			
/i/	/u/	/ei/	/œy/	/au/	/ou/
/I/	/Y/	/ai/			
/ɛ/	/œ/	/ɔ/			
	/a/				

There is no distributional difference between the diphthongs and the monophthongs, both can occur short or long

(according to the rule presented in Chapter II), so they can be said to form an integrated system of steady state and moving, 'dynamic'. (Stebelin-Kamenskij 1960:42) vowels. The quality of the vowels varies a bit according to whether they are long or short (cf. Ch. V, Section 5 and Garnes 1974b). When the nonhigh monophthongs /I/, /Y/, /ε/, /œ/, /ɔ/ and /a/ combine with a following [j], (which can in most of these cases be analysed as deriving from a velar fricative by palatalization), they form in most dialects diphthongal allophones ('combinatorial diphthongs'): [Ij] [Yj] [ɛj] [œj] [ɔj] and [aj]. These diphthongs are entirely predictable from the environment and are therefore not to be considered as phonemes, and few people do indeed consider them as such. This phonemic analysis is isomorphic with the analysis underlying the orthography, each of the phonemes has a separate symbol or digraph:
/i/ : í /I/ : ī /ε/ : ē /Y/ : ý /œ/ : ö, /u/ : ú
/ɔ/ : ǫ, /a/ : a, /ɛi/ : eí /œy/ : œy /au/ : á, /ou/
: ó, /ai/ : æ.

Consonants:

A phonological analysis of the Icelandic consonants poses much more complicated problems than the analysis of the vowels, and no attempt will be made here to solve them all. I will only give a brief summary of the main facts.

a. Stops:

There can be said to be two groups of stops in Modern Icelandic, distinguished in traditional Icelandic grammar

by the terms 'hard' (hörð lokhljóð) and 'soft' (lín lokhljóð). I will follow this tradition and use the orthographic symbols in their phonemic representation. Here we meet the first unclear point: it is disputable whether palatals: [c^h], [j] should be represented as separate phonemes or just allophones of velars. I choose here, more or less arbitrarily, to look on them as allophones of velars. Ignoring the problem of the palatals, then, the stop phonemes can be said to be the following:

Hard: /p/ /t/ /k/

Soft: /b/ /d/ /g/

The hard/soft distinction is very important in Modern Icelandic phonology. In initial position, the hard stops are aspirated in all dialects, and the soft ones unaspirated. Both are voiceless in all environments (and in that respect the notation used above is perhaps slightly misleading). In medial position between vowels, and between a vowel and /j,v,r/, there is a dialect difference. In the North only the hard consonants occur, but in the South only soft ones. Thus taka 'take' is [t^ha:k^ha] in the North, but [t^ha:ka] or [t^ha:ga] in the South. In accordance with this the southern variety is called 'soft speech' (línmáli), and the northern variety is called 'hard speech' (hörðmáli). In Praguean terms the 'hardness' opposition is neutralized in this position, in the North in favor of the hard phoneme, but in the South in favor of the soft one. Following an initial /s/, the hard stops lose their aspiration (or only soft ones occur): spara [spa:ra] 'save'.

This is true of all dialects. In other medial positions the hard/soft distinction can be said to prevail in all dialects. If hard stops follow liquids or nasals, the latter are devoiced: heilt [heilt] 'whole' (neuter) (cf. heil [heisl], *ibid.* feminine), fantur [fantʏr] etc. There are exceptions to this in some northern dialects. When the hard stops are geminated according to the spelling, or they precede /l,n,m/ they are preaspirated (cf. Chapter II, Section 2.): yakka [vəkka] 'walk to and fro' (as opposed to vagga [vəɡːa], [vəkːa] 'a cradle') vakna [vəkna] 'wake up' (intransitive) (as opposed to vagna [vəɡna] 'cart' acc. plur.) Here (as a matter of prejudice as far as this presentation is concerned) I choose to follow the spelling in my phonological analysis, but there are various arguments that can be put forth in support of this (for some, see below and Arnason, forthcoming). I would thus analyse yakka, vagga vakna and vagna phonologically as /vəkka/, /vəɡga/ /vəkna/ and /vəɡna/ and set up a preaspiration rule to derive the appropriate phonetic forms (see p. 37).

b. Fricatives:

The phonological analysis of fricatives is no less problematic than that of the stops. The phonetic forms that occur are:

f	θ	ç	x	s
v	ð	j	ʏ	

Of these /f/, /v/, /j/, /s/ are undisputably phonemic. [θ] and [ð] are in complementary distribution, so there

is good reason to assign them to the same phoneme, call it /θ/ (or, according to the Icelandic spelling, /p/). In the velar and palatal region, there is considerable confusion. It is arguable that [ç] is a separate phoneme from [j], since it is in a minimal opposition with it in pairs like hión [çou:n] 'a married couple' vs. Jón [jou:n] 'John', but some might suggest that [ç] should be analyzed phonologically as /hj/ (see below).

The phonological status of [x] and [ɣ] is not immediately clear. They alternate with each other morphophonemically, and [ɣ] alternates with [ǰ], and [x] alternates with [k^h], and to make things still more complicated, [ɣ] sometimes alternates with [j]. (See Chapter II, Section 3. for some discussion of this.)

c. Nasals and liquids:

The following nasal and liquid sounds are to be found in Icelandic: m, n, ɲ, ŋ, l and r. Of these there are both voiced and voiceless varieties, and it seems most natural to derive the voiceless ones from underlying voiced phonemes. The voiceless varieties are usually predictable from the surroundings, as in the examples mentioned above: voiceless in front of hard stops. (heil [hɛi:l] - heilt [hɛi:t] 'whole', væn [vai:n] - vænt [vaint] 'nice' (fem. vs. neuter)). The only place where there is doubt as to whether voicelessness of a nasal or a liquid is predictable is in initial position in forms like hnota [ɲɔ:t^ha] 'nut' where there is on the surface a minimal opposition with [n] as in nota [nɔ:t^ha] 'use', but

here it can be suggested that the voicelessness derives from an underlying /h/, which can only occur initially. This would, then, be in agreement with the analysis of [ç] as underlying /hj/. I will not commit myself on this issue, but simply ignore it and speak as if it had been decided that the voiceless alternants are merely allophones but not separate phonemes. This does not affect the validity of other comments made here on Modern Icelandic phonology.

In connection with the nasals, there is still a comment to be made on palatals and velars. In general, [ɲ] is confined to palatal environment: lengi [lɛiɲʝi] 'for long', and [ŋ] to velar environment langur [laŋgʏr] 'long' (adj.), whereas [n] appears in other environments. This could be taken as evidence that velar and palatal nasals are allophones of the dental one. However, there are to be found minimal pairs distinguished by a velar vs. dental nasal: lengd [lɛiŋd] 'length' vs leynd [lɛind] 'secrecy', but a case can be made, admittedly on morphological grounds, for an underlying /g/ between the nasal and the dental in lengd: cf. langur [laŋgʏr] 'long'.

d. A final note on [h]:

Asⁱⁿ most other Germanic dialects [h] (largely a devoicing element) occurs in initial position, as in hata [hast^ha] or [ast^ha] 'hate'. It seems reasonable to call this a separate phoneme (with a very defective distribution). If initial [ç], [ʃ], [ʒ], [ŋ] are considered to be derived from /hj/, /hl/ etc., /h/ occurs initially in front

of vowels, liquids, nasals and /j/, but if these segments are separate phonemes, the distribution of /h/ is limited to prevocalic initial position. (See below p.33 on the relation between /h/ and preaspiration.)

Chapter II. MODERN ICELANDIC

1. The length rule

All Icelandic non-compound words are stressed on the first syllable, and a secondary stress is placed on every second syllable, counting from the first one. Examples illustrating this are táka 'take', átla 'be going to' for dissyllabics and álmennak 'calendar', álmennakenna (gen.pl. def.) for polysyllabics. (The acute indicates primary and the grave secondary stress.) I will not formalize here a rule to account for this, but it seems to be quite clear that there is a fairly simple and general one at work¹). My main concern will be the relation of stress to segmental length. Long vowels and consonants only appear in primary stressed syllables.

It is quite generally agreed, I think, that (primarily) stressed syllables are all of the same length. (cf. Benediktsson 1963 and references). Whether they are called structurally long or short is probably of limited consequence, since there is no distinction between short and long syllables in the same environment. They are, however, undoubtedly phonetically longer than unstressed syllables: tala [t^ha:lǽ] 'speak', talar [t^ha:lar] 'speaks', manninn [man:In] 'the man'. Neither long vowels nor long consonants appear in completely unstressed

syllables (though they can be made long under emphatic stress, see Section 4.) The stressed syllables can be said roughly to be of three types:

- (1) (a) VCC... (possibly more than two Cs)
- (b) V:C
- (c) V: $\frac{\#}{V}$

(I am not making any assumptions at this stage as to whether V: should be analysed at some (or all) level(s) as a vowel cluster or whether some CC sequences should be interpreted as C:.) Examples illustrating this are:

- (2) (a) hestur [hɛsɔYr] 'horse'
- vinna [vIn:a] 'work'
- (b) tapa [t^ha:p^ha] 'loose'
- tala [t^ha:la] 'talk'
- (c) ný [ni:] 'new' (fem.)
- te [t^hɛ:] 'tea'
- búa [bu:(w)a] 'live'

The phonological analysis of this data has been a matter of dispute for some time (see e.g. Bergsveinsson 1941, Malone 1952, 1953, Haugen 1958, Benediktsson 1963, Kjartansson 1974, Garnes 1973, 1974a and Árnason 1975). People have looked at pairs like man [ma:n] 'sleeve', 'remember' (1.p. sg. pres.) and mann [man:] 'man' (acc.) and made the observation that the sequences [man] and [ma:n:] are impossible in Modern Icelandic. The fact that there are no stressed VC or V:C sequences has of course led to the conclusion that length cannot be distinctive in both consonants and vowels, because then we would

expect a four way distinction V:C, V:C:, VC and VC: . This can be taken to be generally agreed upon. (Malone 1952 held the view that length was distinctive both in consonants and vowels, but he later abandoned it:1953). What has been a matter of dispute is whether length should be assigned to vowels or consonants on the phonological level or indeed whether length is governed by an underlying feature or features. Benediktsson (1963) sums up the four types of solutions proposed so far for the pair man/mann, as follows:

- (3) (a) /ma:n/ vs. /man:/ (Malone 1952)
- (b) /ma'n/ vs. /man'/ (Malone 1953, Haugen 1958)
- (c) /ma:n/ vs. /man/ (Bergsveinsson 1941, Garnes 1973, 1974a)
- (d) /man/ vs. /man:/ (Benediktsson 1963, Arnason 1975)

I have already mentioned the arguments against solution (a), as its being uneconomical, since no syllables of the form VC or V:C: exist and, as far as I know, no one holds this view any more.

Solution (b), proposed by Malone (1953) and elaborated by Haugen (1958), abstracts quantity from the individual phonemes and makes it a feature of the syllable, the phonetic duration (presumably) being governed by phonological rules operating on the accent marker: /ma'n/ → [ma:n], /man'/ → [man:]. This solution takes care of the above mentioned gap in the distribution of quantity, and connects it with stress²). But apart from setting up

an abstract distinction between placement of stress on the vowel or on the consonant, which is not (or at least not proved to be) directly evidenced by the phonetic data, it runs into difficulty in forms like hestur [hɛsdʏr] 'horse' and sá [sau:] 'saw' or te [tʰɛ:]. If there is an underlying distinction made by the placement of 'accent', we would expect it to appear in hestur and te too. The fact is that the placement of 'accent' is only 'free' in pairs like man/mann. In forms of the type te, ((C)V#) there is no distinction in placement of 'accent'. This difficulty is perhaps not so serious; it could be said to be merely an instance of defective distribution of a phonemic distinction. and, furthermore, this defective distribution is easily explained by the fact that in these forms there is no consonant to place the accent on, so it can only fall on the vowel.

The difficulty is perhaps more serious in forms like hestur. Here, the vowel is always short, so that it cannot be analysed as taking the accent. Maybe this can be explained away as an instance of defective distribution, but/^{the}real problem is that the accent has to fall somewhere, and it never falls on the vowel, so it will have to fall on one of the consonants, say on the /s/. This would give us underlying /hes'tur/. This is, however, not very convincing, since then we would expect the phonological rule mentioned above to give a phonetically long [s:]. But this is not what is generally believed to be the right phonetic form. The s in hestur is phonetically short, or at least

shorter than the s in viss [vis:], which would be correctly predicted by Haugen's analysis /vis'/. I can think of two ways out of this difficulty for proponents of the 'accent analysis'. One is to set up a special rule, which would (at least partially) undo the effect of the lengthening rule in this environment, probably defined as the following consonant. The same result would be obtained by making this environment an exception to the lengthening rule, which would leave a phonological accent with no phonetic consequences; and that is not very convincing. There is perhaps some support for the former escape route in the fact that some people believe that the s in hestur is slightly longer than e.g. the s in ýsa [i:ssa] 'haddock', but this is not accepted by everyone, and we would still have two rules operating in sequence, one lengthening and the other shortening the same segment. The phonetic data are still less compatible with putting the accent on the dental stop in hestur: /hest'ur/, since, as far as I know, no one has ever suggested that it has phonetic length.

The arguments proposed in favour of the (c) solution are based on phonetic evidence in some sense. Sveinn Bergsveinsson (1941, and in a discussion in Benediktsson 1963) resorts to something, which he calls 'Dehnbarkeit' of the segments: "Der Unterschied der zwei Normen (i.e. long: short) bei den Konsonanten ... ist nicht so ausgeprägt wie bei den Vokalen". (1941:84) By 'Dehnbarkeit' I imagine Bergsveinsson means the ability to be lengthened; and as a consequence of the vowels having more 'Dehnbarkeit' than

the consonants, their length vs. shortness is more 'prominent' (ausgeprägt) than that of the consonants. It is not immediately clear what bearing observations of this kind should have on the phonological analysis. I will deal with this question after having described briefly an experiment made by Garnes (1974a), which led her to the acceptance of solution (c). The experiment tested the responses of native Icelandic speakers to synthetic tokens made out of e.g. the sounds corresponding to Icelandic f [f] and s [s]. The length of the segments was varied systematically, and the subjects were told to identify the sound sequences as either is 'ice' (nom.) or iss (gen.). The result of this experiment was that the judgements of the speakers was almost solely based on the length of the vowel (Garnes 1974a:224-269, see also Árnason 1975). The argument is, then, that since speakers use the duration of vowels to distinguish between the stimuli, vowel length must be distinctive and pairs like is : iss; man:mann are to be analysed phonologically as /i:s/ : /is/ : /ma:n/ : /man/ respectively. This would then entail a phonological rule, lengthening consonants after short vowels, and the length of the consonants would then be predictable and not distinctive. Notice that the term distinctive is used in this (somewhat hypothetical) argument for phonemic vowel length in Icelandic in two different senses. In its first occurrence it means roughly: 'used by the hearer to distinguish, or rather try to find out, to what form of his language the noise he hears best

corresponds'. In the hearer's performance of this task, there are quite a few things involved, which are too complicated to go into in detail here; but I shall give a brief description of what I think is going on. The speaker hears a certain noise and he expects this noise to correspond to some forms of his language, in this case the forms is or iss. The noise consists of a complex of sound waves, which is picked up by his auditory organs. The noise is very complex and he has to pick out certain features of this noise, which may help him to decide what to make of it. What would be the best strategy in this situation? Of course to look for the most 'prominent' features of this noise, which would give him clues as to which form it is most likely to represent. (The term 'prominent' is used here in a very vague sense, of course, similar to Bergsveinsson's term 'ausgeprägt'.) I would tend to think that all that the experiment shows is that the hearer uses the vowel sound, which is most 'prominent', to give him clues as to what meaning to assign to the noise he hears. When he hears a long [i:], followed by a certain interval of an g-sound, he judges the noise as being is, since he knows that is should have a long vowel. The fact that this judgement is not invalidated by a following long g may lead people, as it did Garnes, to the conclusion that the function of the long g in the language system is secondary to that of the long vowel. I would think that this is not necessarily the case. It is quite possible that the hearer masters a language system (com-

petence), which uses the consonants to predict phonologically (or systematically) the length of the preceding vowel, even though, when faced with some noise which he is told to fit with his system and interpret according to the rules of that system, he uses the vowel length as a clue to what the noise most likely represents. What I am maintaining is, in fact, a distinction between a language system or competence and its use or the speaker's performance with it. The term distinctive, in its two instances in the argument above, refers in one case to something's being used in the act of interpreting a noise, and in the other case it is used in the structural sense as a linguistic terminus technicus something like: 'not predictable from other phonological features'. This is why the argument is invalid and does not force one to take vowel length to be structurally distinctive.

This discussion of the validity of Garnes' experiment as evidence of phonological structure touches on a very central problem in linguistics. This is that it is not always clear what conclusions can be drawn about 'linguistic structure', whatever its ontological status in fact is, from experimental data involving linguistic behaviour of speakers. This is so, in general, because there are always more things than 'linguistic structure' involved when speakers perform tasks like the one involved in Garnes' experiment. A speaker's knowledge of the structure of his language is not the same as the strategies he uses when making judgements like the ones measured by Garnes.

The fact that it seems to be possible to shunt experimental evidence like this, may lead one to become sceptical of their validity in linguistic argumentation, or perhaps conversely lead one to become sceptical about linguistic arguments, since it would seem to be very difficult to test them with experiments. In fact, I think behavioural experiments and linguistic argumentation are by no means incompatible, but it is very difficult to formulate a linguistic argument and an experiment which could be used to test its validity so that it is clear that the variables of the experiment can be unambiguously related to a unique phenomenon in a proposed underlying 'linguistic structure'.

As must be evident from the preceding discussion, I believe that the best solution is solution (d), which makes vowel length (structurally) dependent on the following consonantism. The arguments in favour of this analysis are presented in Benediktsson 1963 and in somewhat more detail in Árnason 1975. The main argument is morphophonemic in the old sense of the term, and it shows that the morphology becomes hopelessly complicated if any of the above mentioned analyses except for (d) is adopted. Icelandic shows morphophonemic alternation between long and short vowels; we can take the two strong neuter nouns hús 'house' and vor 'spring' as examples:

(4)	nom. sg.	<u>hús</u> [hu:s]	<u>vor</u> [vɔ:r]
	gen. sg.	<u>húss</u> [hus:]	<u>vors</u> [vɔ:rs]

The genitival ending is -g, and when it is added to a stem

ending in a single consonant, the vowel automatically becomes short. If the phonological analysis for the pair hús/húss were /hu:s/:/hus/, as it would be if alternative (c) is adopted, the morphological analysis would work something like this:

- (5) i Length is phonemic in vowels and predictable in consonants
- ii The length of the vowel distinguishes between the nom. and gen. of the word hús
- iii The genitival marker for hús is vowel shortness as opposed to the length in the nominative

But the conclusion iii is intolerable, because, then, if vor and hús are to belong to the same declensional class, the genitive marker for vor would be the shortness of the vowel and the -g in vors would have to be predicted for this form by some strange morphophonemic rule.

Notice that this argument contradicts solution (b) just as strongly as solution (c). In that analysis hús and húss would be phonologically /hu's/ and /hus'/, and the genitive would be marked by the place of the accent, and the same strange morphophonemic rule would have to predict the -g in vors. I think anyone would feel a bit uncomfortable about the following derivation of vors:

(6) Phonological form	/vor'/
(after case has been marked)	
Morphophonemic s-addition	/vor's/
Phonological lengthening	vor:s
Phonetic shortening, etc.	[vɔr·s]

If one believes that phonological solutions should not have intolerable consequences for the morphology, and if one thinks that phonology and morphology are interrelated and should be dealt with with at least intertranslatable theoretical apparatus, one has to accept solution (d). If one does not accept this solution, one will have to deal with length in Icelandic in such a way that the morphological description of e.g. vors will have little or no relation to its phonological description, and this should lead proponents of solution (b) or (c) to believe that morphology and phonology are two fairly unrelated levels of language (or the description of language). This is certainly against the spirit of the age and I think most linguists would find this hard to swallow.

It should be pointed out here, that the solution (d) seems to have a theoretical consequence which some people might not like, namely that a clear distinction is to be made between competence and performance. If solution (d) is adopted, one has to allow for the possibility, in the face of Garnes' experiment, that people use other features than the structurally distinctive/unpredictable ones to distinguish between utterances, when they hear them. (See Jakobson, Fant & Halle (1952:8) for an anecdote concerning the use of the redundant backness of the vowel [ɨ] in Russian to distinguish between two words structurally kept apart by palatal/non-palatal consonants.) I interpret this as evidence in favour of the distinction between competence and performance.

The conclusion of this discussion is, then, that vowel length is context-determined, the context being the following consonantism and, of course, the stress of the syllable. The rule can be stated tentatively like this:

$$(7) \quad \left[\begin{array}{c} \text{v} \\ +\text{stress} \end{array} \right] \rightarrow [+long] / - \left\{ \begin{array}{c} \text{C1} \\ \text{v} \\ \# \end{array} \right\}$$

There are some very important assumptions behind this form of the rule, which are unjustified at this stage, e.g. that length is the marked value (a lengthening rule instead of a shortening rule), and that stress and length are features of the same kind as e.g. dental, consonantal etc. These assumptions involve very important theoretical issues, which I will deal with at different places in this study (Section 4 of this chapter and Chapter V), but for the present purposes the formulation in (7) will do, even though some aspects of it may be questionable. (See Eliasson and La Pelle 1973 for a similar rule for Swedish.)

There is an exception to the length rule, which I should mention now. This is that before sequences, in which the first consonant is a member of the set /p,t,k,s/ and the second of the set /v,j,r/, vowels are long:

- (8) nepja [nɛ:p^hja] 'cold weather'
 vitja [vi:t^hja] 'attend to'
 sækja [sai:c^ha] 'fetch'
 Esja [ɛ:sja] name of a mountain
 (upp)götva [gœ:t^hva] ([gœ:tfa]) 'discover'
 skrökva [sgrœ:k^hva] ([sgrœ:kfa]) 'tell a lie'

depra [dɛ:p^hra] 'sadness'
titra [t^hɪ:t^hra] 'shiver'
sykra [sɪ:k^hra] 'put sugar on'
Esra [ɛ:sra] a man's name
tvisvar [tʃɪ:svar] 'twice'

The sequence /pv/ does not occur, probably prevented by a phonotactic constraint. I will come later to the problem of how to deal with this exception.

Except for the ones just mentioned, then, every cluster in Icelandic has a short vowel in front of it. It is hardly a coincidence, but probably an aspect of the same rule, that gives short vowels in front of long consonants. This fact clearly suggests an analysis of the long consonants as underlying clusters of two identical consonants. This seems to be very plausible. Consider for example the above mentioned pair hús:húss as compared with vör:vörs. The genitive marker is evidently -s, and the difference between the nominative and genitive of hús is best described as the absence vs. presence of a second /s/ : /hus/ /huss/. A low level phonetic rule is perhaps needed to eliminate the boundary between the first and second element of the cluster, if one is conceivable. (see Lehiste 1970: 44). There are of course examples where there is no morphological support for the analysis of long consonants as clusters of two identical ones, for example koss 'a kiss' kunna 'know how', but it seems to be reasonable, in the absence of any evidence to the contrary, to analyse these examples in the same manner: /koss/, /kunna/. It may be

mentioned in passing that this is exactly what Icelandic orthography does, as can be seen from the examples, but the arguments put forth in favor of the present analysis are independent of that.

I can think of one fact which could possibly be considered to contradict the analysis of long consonants as clusters. In words with long consonants in their stems, for example verbs like kyssa 'kiss' and kenna 'teach' or adjectives like viss 'certain', when a consonantal inflectional ending is added to the stem, the consonant loses its length or, in our terms, one of the consonants of the cluster is deleted: kyssti [c^hIsdI] (past, 1., 2. p. sg.) visst [vIst^h] (neuter). This cannot be dealt with simply by a phonological constraint, prohibiting clusters of three consonants in these surroundings, since there are stems, having clusters of non-identical consonants, which keep their clusters intact: herma - hermdi [hærmðI], not *[hærdI] or *[hæmdI]. There are, however, stems with consonant clusters which show behaviour which might be considered to be an aspect of the same phenomenon as appears in kyssa - kyssti and kenna - kenndi. In verbs like verpa 'lay eggs' and adjectives like skarpur 'sharp', the addition of the inflectional morpheme -t causes a considerable weakening of the /p/, or rather its morphophonemic variant [f], so that there is hardly any sign of it, maybe only a slight rounding of the [r]. The following are variant phonetic realisations of verpa in the past tense, verpti and skarpur in the neuter, skarpt: [verfðI], [verfðI],

[væɾ^wɔ̃I]; [sǰarft^h], [sǰar^wt^h], [sǰart^h]. The first instances of each show the effects of a general rule, which turns clusters of two unvoiced stops into clusters of a fricative + a stop, but the other forms show at least a tendency to get rid of the clusters of three consonants by deleting the one in the middle. There is a difference of style between the three forms, the first being the most careful speech, and the others less so. A similar phenomenon is shown by verbs with stems in /-ng/, e.g. hengja [hæɪpʃa] 'hang' (transitive), past: [hæɪndɪ], where there is in most dialects of Modern Icelandic no sign of the stem final /g/, except for the velarisation or palatalization of the nasal. A number of examples of this kind can be cited. If there is a rule eliminating some (but certainly not all) clusters of three consonants, the /sst/ and /nnt/ etc. can easily be included among those.

My conclusion is, then, that Benediktsson's analysis (1963) is the best one so far and that a rule similar to (7) (maybe its inverse, shortening vowels in front of consonant clusters) is operative in Icelandic.

2. Preaspiration

The conclusion of the preceding section was that short vowels are (with the above mentioned exceptions) conditioned by following (underlying) consonant clusters. I would now like to look at some forms which relate to this and may make things more complicated than they seem at first glance, even though I think they can in no way be taken to contradict the analysis proposed here.

I have in mind the so-called preaspiration in forms like hnakkur 'saddle', tappi 'cork', detta 'fall', epli 'apple', kuki 'witchcraft', vopn 'weapon', vakna 'wake up'. There has been some dispute on the phonetic and phonological nature of the sequences represented by the spellings kk, tt, pp, pl, kl, pn, kn. There have been mainly two matters of dispute: the nature of the preaspiration and (in the case of pp, tt, kk) the length of the closure.

Guðfinnsson (1946) would transcribe the forms in question phonetically like this: [n^hak:Yr], [t^ha^hp:I], [d^hε:t:a], [ε^hp:lI], [va^ht.n]. He considers the preaspiration to be, so to speak, the inverse of (post) aspiration of stops in initial position ([t^ha:la]), and takes the closures of the double-written forms to be long and the stop members of the clusters pl, kn, etc. to be 'half-long'. Guðfinnsson does not analyse these forms phonologically, but a phonetic analysis of this sort might lead to a phonological analysis with the preaspiration as a feature belonging to the stop segments in question, (whether it be distinctive or not). Others, e.g. Ófeigsson (1920-1924), Malone (1952) and Einarsson (1927) have transcribed these forms differently, taking the preaspiration to be a separate segment of some sort: [d^hεhta] etc.

Recent phonetic investigations of this phenomenon, Pétursson (1974:188-89) and Garnes (1973), have confirmed that the stop segments in hnakkur etc. are short (i.e. shorter for example than the segments in forms like lögga [lœg:s] 'police') and that the preaspiration takes up a

considerable part of the time of the syllable (word) as a whole. They seem, in other words, to support a phonetic transcription of the type [nahkY_h], [d_hhta] etc., and similarly [ɛhplI], [vahkna]. These facts lead Pétursson (1974: 186) to the conclusion that the preaspiration is merely an instance of the phoneme /h/, also appearing in initial position in forms like hús, hestur etc. The main reason for Pétursson's analysis is that the phonetic properties of the preaspiration are almost exactly the same as those of /h/ (a devoicing of the adjacent voiced (devoiceable) segment: hestur [ɛ_hɛsd_hY_h], hnakkur [n_hak_hY_h]). Pétursson's phonological analysis of the forms in question would then presumably be /hnahkY_h/, /tahpI/, /d_hhta/, /ɛhplI/ and /vahkna/. (It should be noted here that even though the preaspiration behaves phonetically like the phoneme /h/, it does not, within generative phonology, necessarily follow that it is an allophone of /h/.) This kind of phonological analysis does not perhaps have any serious consequences for the rule of vowel length proposed above, since /hp/, /ht/, /hk/ are still clusters in some sense of the term. But h being mainly a devoicing element, affecting an 'adjacent' segment, is a peculiar consonant, and it may not seem very convincing to have it as a part of a vowel shortening (or non-lengthening) cluster. A conceivable solution would be to have the vowel in these forms undergo lengthening and let the preaspiration devoice the second half of the vowel, giving the forms [hnahkY_h] = [n_hak_hY_h], [d_hhta] = [d_hɛhta] etc. and derivations like:

/hnaahkYr/ → hnaahkYr → [naakur]. This could, however, not work for the forms yakna and epli, since there the vowel lengthening rule is not supposed to operate.

An alternative to this is to analyse the preaspirated forms as underlying clusters of two consonants, as in fact is done in the spelling, as a consequence of the fact that the forms had long consonants at some earlier historical stage. There are some problems with this solution, for example a rule is needed to convert underlying /Vpp/, /Vtt/, /Vkk/ into phonetic [Vhp], [Vht], [Vhk] or [V_op], [V_ot], [V_ok]. I don't see at the moment how this rule can be stated naturally within the notational framework of generative phonology, but that in itself of course does not count as evidence against this analysis. There is some evidence supporting this kind of analysis. It is, as in section 1, morphological. Consider the paradigms hvit 'white' (fem.) [kfi:st^h] vs. hvitt (neuter). [kfiht^h] and vis 'certain' (fem.) [vi:s] vs. vist (neuter) [vist^h]. These two pairs demonstrate the morphological formation of the neuter gender in adjectives, which takes place by the addition of the morpheme {-t} to the stem. It is certainly very plausible to distinguish the neuter form of hvit phonologically by an additional phonological segment -t̥ as in the case of vis: vist, giving the phonological forms /kvi:t/:/kvitt/, /vi:s/:/vist/. If we are not allowed to analyse phonetic [ht] as /tt/, we have to look for some other neuter marker in hvit:hvitt, the only alternative being the preaspiration,

however we choose to represent it phonologically. This leaves us in the same impasse as with the genitive marker in Section 1, since we are clearly missing a generalization in the morphology by setting up a special class of adjectives, which take the neuter form by an infixation of preaspiration. In this analysis it would, further, be a coincidence that this same kind of infixation appears in some weak verbs in the past tense and past participles: veita 'offer' [vɛi:t^ha] - veitti [vɛihtɪ] - veitt [vɛiht], whereas other verbs take a suffix starting with a t (lääs - lästi - läst). There is still more evidence to indicate that the preaspiration is not an independent phoneme, but predicted by the following stop. Consider forms like genitive stakks of stakkur 'a kind of coat'. The genitive can be pronounced in two different ways: [sɔhks] or [sɔhxs]. The former is more careful pronunciation than the latter. The variation is determined by the deletion or non-deletion of a morpheme boundary, the deletion allowing a phonological rule eliminating clusters of stops + g to operate. Notice that if this fricativization rule is allowed to operate on the stop, the preaspiration disappears too; as far as I know, the form [sɔhxs] does not exist. If the fricativization rule were to operate on a form like /stahks/ there should be no reason for the preaspiration to be wiped out. The preaspiration could perhaps be gotten rid of by some special mechanism, but if the preaspiration is a phonetic consequence of an underlying cluster starting with a stop, no special equipment of that sort is needed.

Before turning to the phonological rule needed to turn /tt/ etc. into [ht] etc., we will have to look at the forms of the type vatr and epli. As far as I know there are no exact phonetic measurements available for these forms, as there are for the hnakkur-type forms, but it sounds to me that they should be transcribed phonetically as [ɛhplɪ], [vahtn], [vahkna] etc. One could possibly analyse them phonologically like this: /vattn/, /epplɪ/ etc. Then the preaspiration in these forms could be given by the same rule as in the hnakkur-type words. But this solution leads to some uncomfortable consequences in the diachronic and morphological description. There is ample diachronic evidence to show that these words (those of them that aren't recent loan words) have a historically short consonant, e.g. vatr, cf. Engl. water, Goth. wato etc., and there is little reason to believe that they had a long stop in Old Icelandic. This would force us to try to look for a historical change, lengthening the stops in these circumstances. Note that this would seem to be the only place in Icelandic where a consonant lengthening took place and in a rather strange environment for that matter. More difficult than this is synchronic evidence of morphological alternations. Many 'weak' nouns (old n-stems) take gen. plur. with the ending -na: gate-gatna, [g̊a:t^ha] - [g̊ahtna]. Similar alternations occur in the inflection of nouns ending in an elideable vowel + l (cf. Oresnik 1971): nom. ketill (<ketil+r), acc. ketil, [c^hat^hɪl] nom. pl. katlar [k^hahtlar] 'kettle'. These forms show alternations

between [t^h] and [ht], and to keep the preaspiration rule unchanged as /tt/ → [ht] we would have to set up a synchronic rule lengthening the consonant, only to make the preaspiration rule applicable. It seems better to extend the environment of the rule to include the clusters pn, tn, kn and pl, tl, kl.

It now remains to find a natural way of accounting for the process in question. I will not state the rule in full here, but only try to describe roughly what I think is involved. It seems to me likely that we are dealing with anticipated voicing offset of the preceding vowel, and it looks as though this voiceless period takes up some of the time 'allotted' to the consonant. With the (unsatisfactory) notational conventions of generative phonology, the effects of the rule can be described informally in the following way:

$$(9) \quad [+cons] \rightarrow h [+cons] \left/ \begin{array}{l} \underline{v} \\ [+hard] \end{array} \right. \left[\begin{array}{l} +cons \\ +interr \end{array} \right]^{3)}$$

This would change Vtt into Vhtt and Vtn into Vhtn. But this is not enough, since an additional phonetic rule would have to be posited to shorten the stops in htt etc. This may not look very elegant, but if the rule shortening the stops is considered to be a low level phonetic adjustment rule of some sort, this doesn't have to be so bad.

We then reach the preliminary conclusion that the forms hnakkur, tappi, detta, epi, yakna and others of the same sort are to be analysed phonologically like this: /hnakkur/, /tappi/, /detta/, /epi/, /yakna/ etc. The length of the preceding vowel is then determined by the

rule (7) and these forms pose no problem for the analysis of vowel length. Notice that we are forced to order the vowel length rule (7) before the preaspiration rule (9) and its secondary shortening of tt etc., if we want to keep the former as simple as it is in (7). The problem of rule ordering will be touched on in Section 4 of this chapter and I will not go into this here; but notice that the order (9) → (7) is a bleeding order in the sense of Kiparsky (1968a), and should then be marked by Kiparsky's principle, whereas the order (7) → (9) is unmarked by the same principle and should then be the natural order of application.⁴⁾

3. Exceptions to the length rule, syllabification.

3.1

In this section I would like to have a look at the exceptions to the length rule listed under (8) in Section 1. As stated there, vowels are long before sequences of two consonants of which the first is from the set /p,t,k,s/ and the second from the set /v,j,r/. This makes rule (7) inadequate in that it does not account for the length of the vowels in this environment. One has to ask whether there is a natural explanation for this exception, or, to put the question slightly differently, whether it is an exception at all, that is, whether the environment /p,t,k,s/ + /v,j,r/ has something in common with the one already stated in rule (7). If we can find a common denominator of some sort, the rules should be reformulated in terms of that.

It has been suggested by Vennemann (1972:7) and Garnes (1975a:1-3) that this apparent exception to the length rule can be explained in terms of syllable structure. They propose a syllabification which treats the forms neþia, vitia, sakia, Esia, (upp)götva, skrökva, deþra, titra, svkra, Esa and tvísva differently from other forms having intervocalic sequences of more than one consonant. Garnes calls upon a 'sonority hierarchy', proposed by Zwicky (1972), to help to define the environments of a syllabification rule, which gives the desired results. She proposes that the forms with a short vowel be syllabified so that a syllable boundary falls between the two consonants, leaving a consonant following the vowel within the same syllable; the forms with a long vowel she proposes to syllabify so that the syllable boundary falls immediately after the vowel. This would give a syllabification like yak-ka 'walk to and fro', yak-na 'wake up (intr.)' hes-tur 'horse' for the short vowel forms as opposed to ve-kia 'wake up (trans.)' va-kra 'good for riding' (acc. mass. pl.) E-sia etc. for the long vowel forms. Vennemann uses a strength hierarchy with /v,j,r/ as the 'weakest' of Modern Icelandic consonants and /p,t,k,s/ as the strongest to get the same results.

Before looking more closely at the data at hand and these proposals, I would like to make a small digression concerning the 'syllable' in general. It has recently been argued that the syllable should be an essential notion in phonology (see e.g. Fudge 1969, J. Anderson 1969, Fulgram 1970, Vennemann 1972, Hooper 1972, Anderson and Jones 1974,

Basbøll 1974, J. Anderson 1975 and Kahn 1976). It has been shown, e.g. by Hooper, that phonological rules can in a number of cases be stated more naturally if syllable boundaries can be used than if they can't. There seems to be, then, good evidence to the effect that the syllable should be set up as a theoretical construct applicable in the description of languages. It is further claimed by e.g. Pulgram that it is universal and every language uses the syllable as a significant unit in its phonology. In accordance with the claim that the syllable should be incorporated into general phonological theory, attempts have been made to devise rules or principles of syllabification which are (explicitly or implicitly) claimed to have universal application, i.e. every human language is to be syllabified by a universal rule. It should be noted, however, that even though the syllable is a linguistic universal, it does not necessarily follow that there exists a universal rule of syllabification, which can be applied to all human languages to insert syllable boundaries at the right places. It is quite possible that syllabification rules are language specific even though the syllable is a universal unit, just as at least some rules for the expansion of NP are language specific, even though the noun phrase itself is probably universal.

Although it is possible not to believe in a universal syllabification procedure and still believe in the syllable, there is a question connected with syllabification rules, which I think should be dealt with in terms of

universality. The proposals for universal syllabification rules that I know of⁵⁾ are of two kinds. Hooper's syllabification rules (1972) are, typically for Chomsky-Hallean generative phonology, mostly based on the distinctive feature structure of the segments. A typical rule of Hooper's is one which says that a syllable boundary automatically falls between two 'non-sonorant' (i.e. containing the feature [-sonorant]) segments (Hooper 1972: 535). The other approach to syllabification rules is one which bases the rules on the phonotactic structure of the language in question. Pulgram (1970), Anderson and Jones (1974), and J. Anderson (1975) seem to me to represent this point of view. Pulgram's basic principle, for instance, is that all syllables are open, provided this is not prohibited by the phonotactic constraints of the language.

For the moment, I will not try to evaluate the relative merits of these approaches to syllabification; we can call the former the 'segmental approach' (e.g. Hooper's) and the latter (e.g. Pulgram's) the 'phonotactic approach'. (A part of the evaluation will of course be defining the differences between the two, and the difference may, in the final analysis, turn out to be insignificant. More on this later.)

Returning to Modern Icelandic, it is evident that Garnes' proposal (as well as that of Vennemann 1972) is based on the segmental approach, i.e. according to her principle, if a voiceless obstruent is followed by a segment which has sonority (in Zwicky's sense) which is

greater than or equal to that of /r/, then the syllable boundary falls in front of the obstruent, but if the segment following the obstruent has less sonority, the syllable boundary falls between the two segments: va-ka, va-kra, ek-la 'lack', vak-na, vak-ka. The lateral /l/ has less sonority than /r/ in Zwicky's hierarchy, and the breaking point lies, according to Garnes, between these two, as far as Icelandic syllabification is concerned. There are specific problems with this analysis for Icelandic. Apart from the fact that this principle, as it stands, does not take care of forms like biðja 'ask' [bʲiðja], telja 'count' [tʰɛlja], blaðra 'balloon' [blaðra] and gulra 'yellow' (gen.pl.), [gʲylra] which have short vowels and should then be syllabified bið-ja, tel-ja, blað-ra etc., Garnes is forced to set up underlying forms for j and y, which are otherwise unmotivated. She proposes to analyse them as underlying glides /y/ and /w/ in order to put them in the right place in the sonority spectrum, whereas they are usually realized as fricatives, which are less sonorant than l, n and m, according to Zwicky. Even though we grant that Zwicky's sonority hierarchy is valid on a very abstract systematic phonemic level, as Garnes must assume, but which seems to me very dubious, Garnes' proposal has the gross disadvantage of setting up underlying segments which invariably turn up as something else on the surface; since, if we allow this, it will become very hard to find a reasonable way of restricting the form of abstract phonological representations (cf. Kiparsky 1968b).

I will return to Vennemann's proposal later in this section, but first I will consider the problem from the phonotactic point of view. But before I look at the problem in detail I would like to make a few remarks about alternative hypotheses within that general framework. J. Anderson (1975:10) distinguishes between what he calls the 'maximalist' vs. 'minimalist' views. Pulgram's principle mentioned above is what Anderson would call 'final minimalistic', according to which as few segments as are allowed to stand word-finally after the vowel of the syllable are assigned to the coda of each syllable. This would mean, for example, that a form like cider will be syllabified ci-der, since the vowel of the first syllable can stand word-finally without any following consonant. The 'initial minimalistic' view would be to assign as few segments as allowed by the phonotactic rules of the language to the onset of the syllable. Anderson (and implicitly Anderson and Jones 1974) proposes what he calls the 'maximalist view', according to which as many segments as possible (according to the phonotactic principle) are assigned both to the onset and coda of each syllable, and overlap is allowed for. According to this, the form debit would be syllabified like this: $\left[\underset{1}{de} \left[\underset{2}{b} \right] \underset{1}{it} \right]$, where the brackets numbered 1 and 2 mark the limits of the first and the second syllable respectively. In this case the segment /b/ belongs both to the first and the second syllable at the same time. I find the terms 'maximalistic' and 'minimalistic' as Anderson uses them somewhat confusing,

since it seems to me that e.g. 'initial minimalistic' can mean exactly the same as 'final maximalistic'. That is, if you assign as few segments as possible to the onset of the following syllable and there is no overlap and nothing left over, you automatically assign as many segments as possible to the coda of the preceding syllable. Perhaps Anderson's principle can be called the 'overlap principle' to avoid confusion.

It is now interesting to see whether we can use the phonotactic method to give us the syllabification we want for simplifying the length rule in Icelandic. If we start by looking at the examples hestur [hɛstʏr] 'horse' (short vowel) and dysja [dI:sja] 'to bury' (long vowel) we may ask whether e.g. the 'final minimalistic principle' proposed by Fulgram can help us. According to this, as few segments as possible are assigned to the first syllable. We see immediately that this does not work, since /st/ and /sj/ are both permissible word-initially in Icelandic: standa 'stand' and sja 'see', and [ɛ] and [I] can stand finally; [ɛ] in te [t^hɛ:] 'tea' and [I] in the name of the letter i, and then hestur and dysja should both be syllabified in the same way: he-stur and dy-sja. If, on the other hand, we assume that modern Icelandic should be syllabified 'final-maximally', we seem to be getting somewhere. According to this principle, we should assign as many segments as possible to the coda of the first syllable. Then we notice that /st/ is a permissible word-final cluster: hest 'horse' (acc.), ást 'love', etc.

whereas /sj/ is not. This is shown particularly clearly by the inflection of the word dys (f.) 'grave' [dɪ:s], which is derivationally related to the verb dysja 'to bury'. The genitive singular is formed by adding the ending -ar to the stem, and then /j/ appears: dysja [dɪ:sjar]. The /j/ of the stem /dɪsj/ is evidently prevented from appearing in the endless nominative by a phonotactic constraint forbidding the sequence /sj#. According to a final maximalistic principle, then, we get hest-ur vs. dys-ja (the /s/ is assigned to the first syllable in dysja, since /s/ can appear word finally after a vowel). If we look at the other exceptions to the length rule, we see that they will all be syllabified in the same way as dysja by the final maximalistic principle: nep-ja, vit-ja, sak-ja, Es-ja, (upp)göt-va, skrök-va, dep-ra, tit-ra, syk-ra, Es-ra, tvis-var.

But our troubles are not over yet. Let us look at the forms biöja [bɪöja] 'ask', tefja [t^hɛvja] 'delay' and stööva [stœöva] 'stop', which have short first vowels. If we were to syllabify these forms according to the final maximalistic view, we would get biö-ja, tef-ja and stöö-va, since [öj], [vj], and [öv] are not permissible word-final clusters in Icelandic. So we see that these forms, having a short vowel, get the same syllabification as dysja with a long vowel. This indicates that the final maximalistic principle cannot help us to get a syllabification in terms of which we can simplify the length rule.

We have still one alternative within the phonotactic framework, namely Anderson's and Jones' overlap principle. This could perhaps help us to differentiate between dysja etc. on the one hand and biðja, tefja and stöðva on the other. We then notice that [ðj], and [ðv] are impermissible word-initially⁶⁾. According to the overlap principle dysja would be syllabified [dy[s]ja], since /sj/ is permissible word-initially, but biðja and stöðva would be syllabified [bið][ja] and [stöð][va], since [ðj] and [ðv] are neither permissible word-finally nor word-initially, and must then belong to different syllables without any overlap. The fact that /s/ in dysja constitutes an overlap could then perhaps be utilized in the length rule, since evidently single intervocalic consonants will also constitute an overlap between two syllables as in mana [ma:na] 'egg on': [ma[n]a]. But this does not solve our problem either. Forms like venja [vanja], 'habit', temja [t^hemja] 'to domesticate', velja [velja] 'choose' berja [berja] 'hit' with short vowels have intervocalic sequences, which are impermissible word-finally, but permissible word-initially: njóta 'enjoy', ljótur 'ugly', riómi 'cream', and they should then fall in the same category as /sj/ according to the phonotactic overlap principle, i.e. be syllabified [ve[n]ja] etc. with the /n/ etc. forming an overlap exactly like the /s/ in dysja.

3.2

It seems, then, that syllabification according to the phonotactic principle does not help us to find a simple explanation for the exceptional behaviour of the sequences /pj/, /kj/, /pr/, /kr/ etc., with respect to vowel length. This can either mean that the length rule is not to be defined in terms of syllable boundaries or that the phonotactic principle does not work for the syllabification of Icelandic.

At this stage it seems to be advantageous to consider three things. Firstly, it must be kept in mind that all the syllables that we have seen to have length variation in vowels are stressed and it is more than likely that stress and syllable structure are interrelated in some way. If, for example, we assume that stress is assigned before the syllable boundaries that define the domain of the length rule are put in, we could perhaps weaken the claims made by the phonotactic principle, that syllables must not end or start with consonant sequences that don't occur word-finally or word-initially, so as to allow stressed syllables to absorb more consonants than would be predicted by the phonotactic principle.

Secondly, it must be borne in mind that there is an assumption that lies behind the phonotactic principle as applied here, which is perhaps illegitimate. This assumption is that the phonotactic constraints of any language should be defined independently of the syllable and that

the syllable should be delimited in terms of phonotactic constraints, which have in turn been defined independently of the syllable in terms of the feature composition of the segments. We see then that the difference between the segment structure principle and the phonotactic principle is perhaps insignificant, since the phonotactic principle is only one (perhaps illegitimate) step away from the segment structure.

There is a third assumption that lies behind the discussion above, namely that the syllable should be defined derivatively in terms of phonotactics or segment structure. This point of view probably derives from the assumption in generative phonology that the phonological rules apply to syntactic surface structures and that phonology is interpretive. If one believes in generative phonology, one almost has to take this point of view. It is, however, theoretically possible that this view is wrong and that the phonology leads a life independent of the syntax and morphology. This view is represented e.g. by Fudge (1969) and Sampson (1970), who set up a kind of rewrite grammar for phonology with the syllable as the initial symbol or axiom.

In view of these considerations, we will have another look at our problem. Let us assume that we could apply the phonotactic principle successfully to give us some syllabification which can help us to simplify the length rule, and perhaps some other rules. If we then ask ourselves what we have done, the answer will be something

like this: We have used the rules determining permissible and impermissible word-final and word-initial clusters as a heuristic device to show us where syllable boundaries fall. But, evidently, we are then presuming that the phonotactic constraints can somehow be described, and if we ask how, the answer will be that the phonotactics is defined in terms of the phonological segments or phonological features of the language. If we say, for example, that Esja is to be syllabified Es-ja and hestur is to be syllabified hest-ur and we do so because the sequences /sj/ and /st/ show different behaviour with respect to phonotactic constraints in that /st/ can occur word-finally, but /sj/ can not, we are merely pushing the problem away without solving it, because we have not explained why these sequences behave differently in phonotactic rules. It seems, then, that the syllabification problem boils down to the problem of explaining why segments behave differently with respect to phonotactic rules; in our case the problem boils down to explaining the deviant behaviour of /p,t,k,s/ + /v,j,r/, and the only way that can be done seems to be in terms of their phonological properties as segments. This, of course, causes no new problems for us concerning the length rule, since we have already seen that the phonotactic principle does not work. But this may show more generally that the phonotactic principle has little theoretical value, but can only be used as a heuristic device.

Furthermore, the phonotactic principle can only be used as a heuristic device to help us to find plausible places for syllable boundaries, if it is generally true that the same phonotactic constraints are valid in non-final/non-initial syllables as in final and initial ones.

Generative phonotactics of the sort suggested by Fudge and Sampson seems to be in the same boat as an interpretive syllabification theory, as far as the relation of phonotactics and segmental structure is concerned. In generative phonotactics the problem becomes one of explaining why only some combinations of segments (or features) are to be generated in particular languages and more generally why there seem to be constraints on the types of syllables that occur in human languages. In our example, we would have to account naturally for the fact that forms like /dIsj/ are not derived from the initial symbol Icelandic Syllable, and I can see no way of doing so but by referring to phonological features that must somehow be related to phonetic properties of utterances, where the phonological constructs /d/, /I/, /s/ and /j/ are represented.

3.3

Throughout this discussion it has been more or less assumed (without justification) that the length rule in Icelandic is in some way connected with syllable structure, and we have been trying to find ways of de-

scribing the syllable in Icelandic in such a way as to be able to simplify the length rule. It is reasonable to ask at this stage whether this assumption is justified.

We can give at least two reasons for assuming that length has something to do with syllable structure in Icelandic. One is that if the syllable can not be utilized in the description of the length rule, we seem to have to state the exceptions in the rule itself, whereas if we can find a syllabification that distinguishes between the cases where long vowels occur and those where short vowels occur, without making the length rule more complicated, we are evidently gaining something. The question now becomes important, whether there are some other phenomena that we can describe more economically if we adopt some syllabification which we can use in order to make the length rule simple. This reason has to do with our descriptive model, that is, we would like to account for the facts as economically as we can. I will return to this point shortly and try to show that there are other facts which have to be accounted for in the synchronic phonology of Modern Icelandic and which seem to require a syllabification of the same sort as the one we can use to simplify the length rule.

The other reason is more closely related to the data itself and becomes more compelling if we put the matter in historical perspective. As is shown in Chapter IV, Old Icelandic had distinctive vowel length. In connection with this, reference is often made to the rhythmic



structure of the language, and in that context the notion of syllable inevitably forces itself into the discussion. It is often asserted that stressed syllables in Old Icelandic were of three lengths, 'short', 'long' and 'overlong'. (In Chapter IV it is suggested that this variation in length of syllables was the basis for the rhythmic structure of the dróttkvætt- and rimur-metres, and perhaps some other Old Icelandic metres.) But ~~this~~ changed in the history of Icelandic in that 'short syllables' were eliminated by lengthening the vowel and the 'overlong' type was eliminated by a vowel-shortening. This brought about a drastic change in the prosodic structure of the language. It can be said that all stressed 'syllables' in Modern Icelandic are long, because if the vowel is short, it is always followed by a long consonant or two or more consonants. It seems to me that a description both of the historical changes and the synchronic facts of Old and Modern Icelandic that makes no mention of these facts is defective, and a description of these facts without the notion of syllable will always be at least very clumsy, if not factually and theoretically wrong.

If we don't want to or are not allowed to use the syllable, we could for example describe the above mentioned facts something like this: In Old Icelandic a stressed vowel was either short or long. In the history of Icelandic, a short, stressed vowel became long if it was followed by one or no consonant (and /p,t,k,s/ +

/v,j,r/). Among other historical changes that hit Icelandic was a shortening of long, stressed vowels, if two or more consonants (except for /p,t,k,s/ + /v,j,r/) followed. This resulted in the length of stressed vowels becoming completely determined by context. If we are not allowed to use the syllable, this is as far as we can go. In the description given here, apart from stating that, perhaps accidentally, these changes resulted in the loss of length as a distinctive feature in the Icelandic vowel system, no attempt is made to relate one change to the other. It seems in these terms to be a mere coincidence that both of these changes occurred. But it is hardly a coincidence, and this has always been assumed by lumping the two changes together and calling them the 'quantity shift'. But what is the justification for lumping these changes together? The answer is that they seem to have a common aim, namely to make all stressed syllables of the same length. (This is true at least for forms which have no more than two consonants following the vowel. The claim is not as convincing when forms with more than two consonants following the vowel, like fisks, the genitive of fiskur 'fish', are taken into account, but there is a strong tendency to simplify many of these clusters of three consonants, cf, e.g. brjósts [brjous:] the genitive of brjóst 'breast' and volgt [völt], the neuter of volgur 'warm'.) It seems, then, that the syllable played a major role in the development of length in

Icelandic, and to avoid this concept in dealing with quantity in Modern Icelandic seems to me to be wrong.

In accordance with the belief just stated, that the syllable is closely connected with the length rule, it is natural to hypothesize that the domain of the length rule is the syllable. This is not necessarily true, but it seems at least a quite plausible working hypothesis, and in what follows I will assume that this is so.

But as we have discovered, it is no simple task to discover what this 'domain' of the length rule is. We have already seen that there are many theoretically possible ways of syllabifying bisyllabic forms with intervocalic sequences of more than one consonant. In forms like hestur, Esja and biðja there are, mathematically, three options available for each form: he-stur, E-sja, bi-ðja; hes-tur, Es-ja, bið-ja; hest-ur, Esi-a, biðj-a, not to mention the alternatives that become available when we allow ^{for} overlap. Our task is to select one of these options for each form and to justify our selection with some data or theoretical arguments. One of the things we can use, heuristically at least, to try to decide which option to choose is, of course, the length rule itself. We will evidently want to ask which option for syllabification can best help us to simplify the length rule. Someone might object to this and say that the argument is circular: You are looking for the unknown thing A (the length rule) and the un-

known thing B (the syllable), but you are using B to look for A and A to look for B, but you know neither what A nor B is, so how can you use them? But this positivistic objection is not so serious, since a search like the one we are undertaking does not have to adhere to the rules of logical proof; one is allowed to set forth hypotheses about things that one does not know and then see whether they fit the things we can observe. But it must of course be borne in mind that even though a hypothesis fits a certain bit of data, it is not necessarily the only right one, and there may be more than one hypothesis that fits particular pieces of data, and the fact that a hypothesis fits does not prove anything about its correctness.

With respect to the simplicity of the length rule, a syllabification like hest-ur, bió-i-a and Es-ia seems to be optimal. If the length rule were to operate on forms like these, it could be stated simply as lengthening vowels in stressed syllables that end in no more than one consonant and/or shortening vowels in syllables that end in two or more consonants. We can state it like this:

$$(10) \quad V \longrightarrow V / \text{--- } C^1\$ \\ \text{and/or } V \longrightarrow V / \text{--- } C_2\$ \\ (\$ = \text{syllable boundary})$$

Let us then tentatively suggest that the forms are to be syllabified in the way described above. The principle would be that two consonants following a

stressed vowel are assigned to the preceding syllable, except when the two consonants in question are /p,t,k,s/ + /v,j,r/, in which case only the first member of the cluster is assigned to the preceding syllable. If only one consonant follows, it is by the same token assigned to the preceding syllable. This syllabification can perhaps be called final-maximalistic in some sense, since as many consonants as allowed by some constraint are assigned to the coda of the syllable.

The main advantage of this syllabification (if it can be called syllabification at all) is that the environment of the length rule, if defined in this way, will be exactly the same in monosyllables and polysyllables, whereas if we were to adopt the syllabification suggested by Garnes (1975a) and Vennemann (1972), where the syllable boundary falls before one intervocalic consonant (as well as /p,t,k,s/ + /v,j,r/) the length rule can not be stated as simply. In the latter case the rule will have to be in two parts. One part would account for monosyllabic forms like hest, nes 'cape', where the vowel is short if followed by two or more consonants, but long if followed by one or no consonant within the same syllable. Another part of the length rule will have to account for the length in polysyllabic forms, where the vowel is short if one (or more) co-syllabic consonant(s) follow(s), but long if the syllable is open, i.e. if no consonant follows within the same syllable.

The most obvious disadvantage of the syllabification suggested here is that it does not follow the 'law of finals' (cf. e.g. Vennemann 1972 and J. Anderson 1975), since the syllabification of grenja and biöja (biöj-a, grenj-a) gives syllables that end in clusters that are impermissible word-finally in the language. And if this law (as well as the law of initials) is a universal constraint, this syllabification should of course be viewed with scepticism.⁷⁾

A weaker claim would be that what we are suggesting is not syllabification, but merely a delimitation of the domain of the length rule. But then, of course, we will have to ask ourselves what exactly this unit is. If it is not a syllable, then what is it? The question immediately turns into the one, whether this unit can have some other function in the phonology of Modern Icelandic, whether, for example, some other rules seem to be defined in terms of it. If this turns out to be the case, we may feel justified in giving this unit a major status in the phonology of Icelandic. We will therefore investigate whether there are other things in the phonology of Modern Icelandic which would become more easily explainable in terms of a syllabification of the sort suggested above.

In forms like hestur, grenja, biöja etc. the consonant following the (short) vowel is often said to be half-long (cf. Guðfinnsson 1948:68-69 and Ófeigsson 1920-24:XVIII-XIX): [hæs.tYr] [græn-ja], [biö-je].

There seems to be some justification for this. I have, for example, made spectrograms of my own speech in the utterance hessi hestur, which show that the [s] in hestur is considerably longer than e.g. the [θ] in hessi, or for that matter longer than any other consonant segment in the utterance except the long [s:] in hessi, which is again noticeably longer than the one in hestur. Although the term half-long seems to be quite appropriate for this phenomenon on the evidence mentioned above, I am not sure that there are not other features as well that characterize consonants in these environments; one should perhaps use some more meaningless term, 'tense' for example. If we now have a look at the distribution of this phenomenon, we see that it must be predicted by things similar to those that affect the length rule. The consonants are half-long or 'tense', if they follow a short vowel and precede a consonant. The distribution of this 'tenseness' is independent of whether the word is monosyllabic or polysyllabic, that is, we have hest [hæs·t], ^{vans} / [væn·s] and ^{glæðs} / [glæð·s] with 'tense' consonants just as in hestur, gleðja and grenja. It seems very tempting to try to explain the distribution of these tense consonants in terms of their place in the syllable (cf. Hoard 1971) but in that case we meet again the fact that this 'tensing' takes place regardless of whether the consonant appears in a monosyllable or in a polysyllabic word. This makes it impossible to capture this phenomenon in

a simple way with a syllabification like the one proposed by Vennemann and Garnes. Not only are the consonants tense before another consonant, regardless of whether that other consonant is word-final or not, which would make the environment for a conceivable 'tensing' rule defined in terms of Garnes/Vennemann syllabification vary according to whether the forms were monosyllabic or polysyllabic, but also the [s] in forms like nes [nɛ:s] would be in the same syllabic environment as the [s] in hestur (both closing a syllable), but with a difference in tenseness. This shows us that if the tenseness of consonants has something to do with syllable structure, this syllable structure cannot be the one proposed by Vennemann and Garnes. But if we adopt the syllabification proposed here, the environment will always be the same, namely, when a post-vocalic consonant is followed by another consonant within the same syllable, the former is 'tensed'.

But again we have no guarantee that this 'tenseness' of consonants has anything to do with syllabification. Furthermore, even though we grant that it has something to do with syllabification, the argument is rather weak as independent evidence for the syllabification we are proposing, exactly because the distribution of tense consonants seems to be related to the distribution of short vowels. It is quite conceivable that the 'tenseness' of the consonant is governed by the same general rule as assigns length and shortness to vowels. We have to admit, then, that the 'tenseness' of consonants

is not a very strong independent argument for the syllabification suggested above.

Another feature could perhaps be taken as evidence for the syllabification proposed here, with as many consonants as possible belonging to the first syllable. This argument is hardly very strong either, but I will mention it anyway. As is shown in Section 2 of this chapter, the 'hard' stops /p,t,k/ are preaspirated when geminated or in front of l, n and m. Here, the preceding vowel is always short: epli [ɛhplɪ], vatn [vahtn], rytmi [rɪhtmɪ] 'rhythm' etc. If we look at the environment where preaspiration occurs, we see that it occurs independently of whether the clusters are word-final or intervocalic. In this respect the preaspiration shows the same behaviour as the length rule and 'tensing' of consonants, and it evidently can't be sensitive to a syllabification like the one suggested by Vennemann and Garnes, since then, in monosyllables the preaspiration would occur on p, t, k if followed by another cosyllabic p, t, k, l, n, m, but in polysyllabics it would be triggered by a heterosyllabic consonant following.

In this case, as in the others, it is of course possible that the preaspiration rule is not sensitive to syllabification at all, but if syllabification has something to do with the preaspiration, it must be one that treats /vatn-/ in vatnið 'the water' and vatn 'water' in the same way in both cases, and our 'maximalistic' syllabification is such a syllabification. As a weak

argument for the case that the preaspiration has got something to do with syllabification I can cite compounds like litlaus [li:tloey:s] 'colourless' (from litur 'colour' and laus 'free (of), without') and saknemur [sa:knai:mYr] 'peccable' (from sök (gen. sakar) 'guilt' and memur 'susceptible (to)'). In these forms, even though the /t/ and /k/ precede /l/ and /n/ respectively, the preaspiration does not occur. This is evidently because there is an internal word-boundary between the two parts of the compounds. This internal boundary must imply a syllable boundary, and if we state the preaspiration rule so that it can't apply across a syllable boundary, these exceptions are automatically accounted for. But the trouble is that there are other ways of explaining why the preaspiration does not occur. We notice that the morphs lit- and sak-, and also the second parts of the compounds, have long vowels. This can be taken to show that the constituents are semi-independent words that have gone through all phonological processes, including the length rule and the preaspiration rule (which does not have any effect on the forms lit- and sak-), before being amalgamated into compound words by some special weakening of the word-boundary. In that case the forms lit- and sak- become just regular monosyllables. (I will return to compounds shortly.)

If we try to sum up what we have said in this section about syllabification, we can say that there is some evidence that the length rule is connected with

syllabification. We have also seen some facts that seem to favour the hypothesis that this syllabification is 'final maximalistic' in a special sense rather than, for example, of the sort proposed by Garnes and Vennemann. The syllabification we have proposed can be described in the following way:

- (11) If a stressed vowel is followed by one consonant, assign the consonant to the preceding syllable, and if the vowel is followed by two consonants, assign both to the preceding syllable, except if the first consonant is one of the set /p,t,k,s/ and the second of the set /v,j,r/. In the latter case the syllable boundary is to be set between the two consonants.⁸⁾

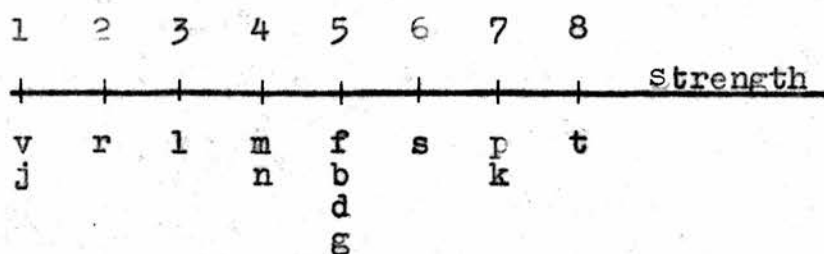
The grounds for adopting this syllabification are, admittedly, weak, but there seem to be no strong internal arguments against it either, and if this syllabification is adopted, the length rule will be as simple as can be.

3.4

I have suggested a syllabification which makes the length rule very simple. Apart from the question whether this syllabification should be preferred to some other syllabification, for example the one proposed by Garnes and Vennemann, there is another question left unanswered. This is the question why the sequences /p,t,k,s/

+ /v,j,r/ behave differently from other intervocalic sequences of two consonants. This can be dealt with independently of which of the alternative ways of syllabifying is adopted, since in any case /p,t,k,s/ + /v,j,r/ will be exceptional.

As mentioned above, Garnes proposes an explanation in terms of a sonority hierarchy. We saw that this explanation is problematic, since it entails underlying forms for /v/ and /j/ that seem otherwise unmotivated. Vennemann proposes a hierarchy of a slightly different sort, which he defines in terms of what he calls the strength of Modern Icelandic consonants. He proposes a tentative scale of consonant strength as follows (Vennemann 1972:6):



(This scale is evidently incomplete, since it does not mention the fricatives [θ], [ð] and [ɣ].)

As can be seen from the diagram, Vennemann considers /v/, /j/ and /r/ to be the 'weakest' of Modern Icelandic consonants. It is not self-evident what can be meant by weakness or strength of consonants. Vennemann wants to explain what he means on phonological grounds and for example cites as evidence for the weakness of /r/ that it seems to be more susceptible to assimilation to or reduction by following consonants.

He mentions that /r/ is always (i.e. in all dialects) devoiced in front of /p/, /t/ and /k/, but /l/, /n/ and /m/ only sometimes (i.e. in some dialects). He also mentions that only /r/ is weakened or lost before /s/ and some other consonants or consonant sequences (but e.g. /n/, /l/, /m/ remain). Vennemann is probably referring to pronunciations like [væstfɪskYr̥] of orthographic vestfirskur 'from the Western fjords', more formal [væstfɪrskYr̥]. It is not true that only /r/ disappears or weakens in this environment, because pronunciations like [ɛ̃skYr̥] for orthographic enskur 'English', more formal [ɛ̃nskYr̥] are quite frequent. (Perhaps someone would maintain that the nasalization left on the vowel is a sign of the greater strength of /n/ than /r/, but of course there are no signs of a 'rhotization' caused by /r/ anywhere else in the phonology, whereas nasalization is a very natural process in Icelandic, and probably in any language. So the /r/ does not have as good a chance of leaving any trace after it when it disappears as /n/ does.)

As a sign of the weakness of /j/ and /v/, Vennemann mentions that they tend not to occur in front of [i] and [u] respectively. But this could just as well be caused by a phonotactic restriction that forbids a sequence of two segments that are too much alike. A restriction of the same type probably forbids sequences of a stop plus a homorganic nasal in initial position: /pm/, /tn/, /kn/ are not allowed word-initially in Icelandic,

and similarly a stop plus a homorganic fricative does not occur word initially. In general, the idea of a strength hierarchy in consonants, though interesting, seems to be very difficult to substantiate.

One way of trying to substantiate a claim of a strength hierarchy is probably to investigate the phonotactic behaviour of the segments (cf. Sigurd 1955). It is, for example, interesting to note that initial clusters of /j/ + /l/, /n/ or /m/ don't exist (*/jl/, */jn/, */jm/), whereas /j/ following /l/, /n/, /m/ is natural: ljótur, 'ugly', njóta 'enjoy', mjólk 'milk'. This could perhaps be interpreted as showing that /j/ has a tendency to stand nearer the vowel than /l,n,m/, when cooccurring with them. Similarly, final clusters like /lr/, /nr/, /mr/ hardly occur (/mr/ occurs word-finally in forms like kumr, which are derived secondarily from verbs like kumra, forming an action name of the same meaning), whereas /rl/, /rn/, /rm/ are regular: harm, 'sorrow' (acc.), Karl, 'Charles', barn, 'child'. (It must be admitted that in most dialects of Icelandic the clusters /rn/ and /rl/ have become [rd̥n] and [rd̥l] or [dn] and [dl] respectively, but there are still some dialects which retain the older pronunciation [rn] and [rl]). It seems then, that /r/ tends to stand closer to the vowel than /l/, /n/ or /m/.

In the light of this, one could perhaps make the generalization that /r/ and /j/ have a greater tendency than /l/, /n/ and /m/ to stand near the vowel

nucleus, and if this holds, one can call that a weakness (or more sonority, cf. Hooper 1976) of /r,j/ compared to /l/, /n/, /m/. I have not been able to find similar arguments for the 'weakness' of /v/.

If one could establish a hierarchy among the consonants either along the lines suggested here or along the lines suggested by Vennemann, or both, the fact that /p,t,k,s/ + /v,j,r/ show exceptional behaviour as an intervocalic sequence could perhaps be explained as some sort of a consequence of their being on opposite ends of a strength scale. This can perhaps be made more plausible if we say that the tendency of /j/, /r/ (and hopefully /v/ too) to stand next to the vowel in a syllable forbids forms like vekja and Esja being syllabified in a way that would leave a segment of the 'strongest' type between it and the vowel nucleus, so only one consonant is assigned to the first syllable. The phonemes /l/, /n/, /m/ show more independence and allow /p,t,k,s/ to stand between them and the vowel nucleus, and the forms vakna, ekla etc. are syllabified vakn-a, ekl-a etc, and this is why they have preaspiration and a short vowel. The fact that venja and biōja, stōōva, viōra etc. have short vowels can be explained as a consequence of the fact that /n/ and [ō] (and presumably /m/, /v/, /l/ etc. too) do not have so much strength as to forbid a syllabification venj-a, biōj-a, stōōv-a, that is /j/, /v/ and /r/ can tolerate them between themselves and the nuclear vowel.

If this bears some relation to the facts, we can restate the syllabification principle (11) and make it look more natural. If we say that /p,t,k,s/ have a strength equal to or greater than the index i and /v,j,r/ have strength less than or equal to the index i, we can restate the principle in the following way:

- (12) If a vowel bearing primary stress is followed by one consonant, assign the consonant to the preceding syllable, and if the vowel is followed by two consonants, assign both to the preceding syllable, if it does not result in a coda where a consonant of strength greater than or equal to the index i intervenes between the nuclear vowel and a consonant of strength less than or equal to the index i.

Another possible way to explain the exceptional behaviour of /p,t,k,s/ + /v,j,r/ is to look for segmental features in these segments, which could be used in a syllabification rule. In order to do this properly, one would, of course, have to set up a distinctive feature system for the Icelandic consonant system as a whole, and it would go beyond the scope of this study to do so. I would, however, like to mention very briefly some facts that may indicate that a solution along these lines is also possible. The central question is whether we can make /p,t,k,s/ and /v,j,r/ form natural classes of some sort. If we can, for example, justify some common

feature or features for /v,j,r/ on independent grounds we will feel confident that they form a natural class. As mentioned above, /j/ and /v/ are phonetically most like voiced fricatives. Admittedly, they sometimes can be said to be realized as approximants (with open approximation), but they do so no more than the other voiced fricatives [ð] and [ɣ]. They are, however, unlike [ð] and [ɣ] in that they occur initially, whereas [ð] and [ɣ] don't. The initial counterpart of [ð] is voiceless [θ], and no velar fricative occurs initially in most varieties of Icelandic. In initial position, voiced and voiceless labiodental fricatives, /f/ and /v/, are in opposition: vara [va:ra] 'last' vs. fara [fa:ra] 'go', and similarly initial [j] and [ç] distinguish between minimal pairs: Jón [jou:n] 'John' vs. hjón [çou:n] 'married couple'. In the latter case it is possible to analyse [ç] as underlying /hj/ so that it is not certain whether the voiced and voiceless palatal fricatives should be taken as two underlying phonemes (cf. Chapter I, Section 3). However that may be, the fact remains that /j/ and /v/ are the only voiced fricatives that are in opposition to other phonemes in initial position. This may perhaps be taken as evidence that /v/ and /j/ are the only underlying voiced fricatives. [ð] can be said to be a voiced allophone of the phoneme /θ/ (orthographic þ), since [θ] and [ð] are in complementary distribution. It is not clear how [ɣ], orthographically g, should be analysed phonologically.

It alternates morphologically both with a voiceless fricative as in dagur [d̥a:ɣr] 'day' dags [d̥axs] (idem, gen.) and with a stop [g̊] as in saga [sa:ɣa] 'story': sagna [səɣna] (idem gen. pl.). It is not obvious what underlying form is to be chosen if one systematic phoneme is to represent all three variants, but obviously one candidate will be a voiced fricative. In that case, /v, j, ɣ/ would form a class of voiced fricatives. But /ɣ/ would behave differently from the others since it would not have a voiced fricative allophone initially. So there seems to be at least some reason to keep [ɣ] apart from /v/ and /j/.

To touch briefly on the question of what to do with /r/, it seems obvious that it can be called continuant in Chomsky and Halle's terms (Chomsky and Halle 1968:318). The fricatives /j/ and /v/ must obviously have this feature too. /r/ has two allophones, a voiced one and a voiceless one, and if we can say that voiced is the unmarked value of the feature of voice, when appearing in /r/, it will have in its fully specified underlying representation the features [+continuant, +voiced]. The sonorants other than /r/, namely /l, m, n/ will presumably be [-continuant]. It seems fairly clear, then, if we ignore the unsettled question of [ɣ], that /j, v, r/ could be distinguished from other Icelandic consonants as voiced continuants, and on the basis of that classification and some feature (or features) making /p, t, k, s/ a natural class, which should not be

difficult to find, it should be possible to define the restriction on the syllabification rule. I will not investigate this alternative further here, since so many questions concerning the whole phonology of Modern Icelandic immediately arise, and an account along these lines will probably only be possible within a more or less complete model of the phonology of Modern Icelandic.

4. Stress

In our discussion up to now, a very important question has been dodged. This is the question of how stress relates to the whole business of length and syllabification. In dealing with syllabification and stress, however, it has been assumed, sometimes tacitly, that syllabification and length are dependent on stress in Icelandic. In terms of ordering this means that stress is distributed before the syllabification (12) takes place and the length rule operates. In this section I would like to try to justify this assumption, in part at least.

As briefly mentioned at the beginning of this Chapter, the main rule of word stress is that noncompound words have primary stress on the first syllable. There was also said to be a secondary stress on the third syllable and every second syllable after that. These rules are manifested in the following way (the number 1 above a syllable represents primary stress and the number 2 represents secondary stress, and if no

number appears above a syllable it means that that syllable is completely unstressed): take¹ 'take', ætla¹ 'be going to', almanak¹ ² 'calendar', almanakanna¹ ² ² (idem gen. pl. definite). This simple stress pattern is disrupted in compound words. The structure of compound words in Icelandic is rather complicated, and there is no room to investigate that matter in any detail here, but a few superficial remarks are needed. It seems to me that there are many types or degrees of compounds in Icelandic. There are cases where there is doubt whether the forms in question are to be considered compounds or not from the phonological point of view. As examples of this type we can cite forms like vitlaus 'foolish' (literally 'wit-less') [vihtlœy:s] and torfara 'obstacle' [t^hɔrfai:ra] (derived from tor-, a prefix, signifying difficulty and fara a root related to the verb fara 'go'). As a sign of the non-compound nature of these forms we can refer to the fact that they have short first vowels even though their first constituents end in single consonants: vit- and tor-. The final consonants of these first parts also undergo phonological rules that operate within simple words. The final /t/ of vit- is preaspirated in front of the /l/ and the /r/ of tor- is devoiced by the following voiceless /f/. Even though the form torfara shows non-compound-like behaviour in the shortness of the vowel and the voicelessness of the /r/, it is in at least one respect compound-like as far as segmental phonology is concerned.

This is that it has a voiceless [f] medially, which normally occurs only initially, its voiced counterpart, [v], occurring medially. Both of the forms show signs of being compounds rather than non-compounds in that the second components bear secondary stress and have long vowels. Forms, similar to the ones mentioned above, but with a looser connection between the two parts, are litlaus 'colourless' [lɪ:tlœy:s] and torlevstur 'difficult to solve' [t^hɔ:rlœistYr]. The form litlaus has a long vowel in the first component and no preaspiration on the /t/, and the form torlevstur has a long vowel in the first component. (The latter word can also be pronounced with a short first vowel, in which case it shows similar behaviour to torfara except for the fact that the sequence /rl/ does not become [rɫ] as it usually does in noncompound words.) The forms litlaus and torlevstur are definitely phonological compounds, since there are rules that are blocked by the existence of some kind of boundary between the two parts (See Orešnik 1971 for an enumeration of such rules). What this boundary actually is I will leave open for the moment, but it seems not unlikely that we are dealing with some kind of a weakened word boundary. The examples given above seem to indicate that there is a 'cline of closeness' of connection between constituents of compound words in Icelandic. The 'closest' compounds mentioned here are, then, vitlaus and torfara and the 'loosest' are litlaus

and torleystur pronounced with a long first vowel, and in between is torleystur pronounced with a short vowel.

The problem of compound words is related to the problem of stress in the way that, at least for the 'looser compounds', the rule for the distribution of secondary stress mentioned above is broken. Forms like ¹ litlaus ² have a stress on the second syllable, since that is the first (and in this case the only) syllable of the second part of the compound. Similarly, varð- ¹ hundur ² 'guard dog' (varð- is a stem meaning 'guard', cf. vörður 'a guard'; hundur means 'dog') has a secondary stress on the first syllable of hundur rather than on the third syllable of the word as a whole.

There is probably more than one way of accounting for these facts. One could for example say that Icelandic has one stress rule something like this:

- (13) Primary stress falls on the first syllable of every word and a secondary stress falls on every second syllable, counting from there, except when the word is a compound. If the word is a compound, then a secondary stress falls on the first syllable of every new constituent of the compound.

This would be an incomplete formulation of the rule. It would for example have to be expanded in order to take care of forms like rakameistari 'a qualified barber'. This is a compound consisting of two trisyllabic forms, which take a secondary stress on the third syllable

when standing as independent words: ¹ rakara and ¹ meistari.

In rakarameistari the strongest stress is on the first syllable, and on the first syllable of the second part of the compound is another fairly strong stress peak, and on the third and the sixth syllable there are weak stresses, which are reflexes of the secondary stress which appears on the component forms when they stand independently. If we use the numbers 1,2,3 to indicate relative strength of stress, the stress pattern of the compound can be indicated like this: ¹ rakara ³ meistari. Although seemingly complicated, an account along these lines seems at least conceivable.

An alternative way of accounting for the stress pattern of Icelandic is to make use of the transformational cycle as Chomsky and Halle do in dealing with stress in English (Chomsky and Halle 1968), and I propose that this is a better way of dealing with the phenomena. In this case, stress would be assigned at two levels, the noncompound level and the compound level. In order for this to work, I have^{to}/set up a mechanism of some sort for generating compounds. I propose that this mechanism can be described as a rule weakening the boundary between the two (or more) lexical items to be bound together as a compound word. This would mean that in their most abstract forms compound words are such that they have between their constituents full word boundaries. The components are assigned stress in the regular way as if they were independent words.

And if the boundary between two words has been weakened by the compound-forming rule, the stress pattern is readjusted, making the first part of the compound the strongest and at the same time making all other stress peaks comparatively weaker. I will not commit myself yet as to what this second stress rule will look like; it may be that it will just be the main stress rule reapplied on a later cycle, the units now being components of the compounds rather than syllables. (This can perhaps be justified by expressions like andskotans-diöf-ulsins-helvítis asni 'bloody-fucking-blooming fool', where the sequence preceding the noun asni 'fool' is a complex of three swearwords. In my speech anyway the third part of this compound swearword has more stress than the second and the first has the strongest stress. Perhaps this reflects the same rhythm as appears in non-compound words, that is, an alternation between stressed and an unstressed unit.) Within this framework the derivation of rakameistari could be something like the following. (## here denotes a full word-boundary and # a weakened word boundary, and the numbers indicate relative strength of stress.):

Main stress rule:	1	2	1	2
	rakare##	meistari		
Weakening of word-boundary:	1	2	1	2
	rakare#	meistari		
Readjustment of stress:	1	3	2	3
	rakare#	meistari		

(It may be unfortunate to use the same type of notation for the secondary stress assigned by the main stress

rule and the one deriving from weakening of the primary stress, since they are definitely not the same phenomenon, but in an informal presentation like this one, I hope it does not matter.)

I need hardly point out that there are many loose ends and unanswered questions still to be dealt with, and this may not even work in principle. It is not clear to me, for example, what governs the application of the rule weakening the word-boundary. It seems doubtful that all compounds can be listed in the lexicon, since the process of forming compound nouns and adjectives of the sort described above seems to be very productive; one can make them up on the spot, so to speak, when the need arises. I can easily form new compounds like þorskastríðshetja 'cod war hero' and Edinborgar-stúdent 'Edinburgh student' etc. There seem to be more restrictions on forming new verbs by such a process. Another problem is mentioned above, namely that different degrees of closeness of the compounds seem to be allowed for; in our terms, the weakening of the word-boundary seems to be allowed to be of different degrees. Forms like vitlaus and torfara seem to have a very weak internal boundary, and the form torleystur with the alternative pronunciations with a long or a short vowel in the first component seems to have varying degrees of closeness of connection between the two components. Evidently, the whole problem of compounds in modern Icelandic is too complicated to be solved here, but I think that what we

have seen so far is enough for us to base on it an argument for stress being distributed independently of length and for the length rule referring to stress. The argument is admittedly rather tenuous and complicated, but in the absence of any evidence to the contrary it should, I think, be considered.

Let us take as examples two compound words which show 'considerable closeness of connection' (or weakness of the internal boundary) between the two constituents: þjóðvísa 'folk song' [θjoudvi:sa] and leikvöllur 'playground' [lei:kvœdlyr]. The former has the constituents þjóð- 'people' and vísa, 'a verse, a song', and the latter has the constituents leik- 'game, play' and völlur 'a field'. From these forms we can observe two things about the length rule.

Firstly we see that the first component of þjóðvísa has a short vowel in a stem which ends in a single consonant. This means that the following /v/ is included in the environment of the length rule when it operates on the compound. This must be taken to mean that the length rule is applied after the compound is formed by reduction of the word-boundary. And if we assume, as above, that the main stress rule applies before this reduction, it follows that it also precedes the length rule. (I must stress that it is by no means necessary that this order be extrinsic. The length rule will have to be defined so that it only applies to stressed syllables i.e. it is not applicable unless stress has already

been assigned, and it can be made a recurrent rule, which automatically applies whenever its structural condition is met). In the form leikvöllur, the vowel of the first component is long, and this is because the consonant sequence following is /kv/, a long vowel environment.

Secondly we see that the second constituents of the compounds also behave according to the length rule: we have a long vowel in -vísa and a short vowel in -völlur. That the actual duration of the vowel in vísa in the compound is the same as in every instance when it occurs as a separate word, is not true, and the average duration of the vowel in vísa as a second part of compounds is probably less than the average duration when vísa is a separate word, but we are not talking about absolute duration, but structural length. We can see that there is a phonological difference between a long and a short vowel in the second part of compounds like hióóvísa, by comparing the nominative singular with the genitive plural hióóvíсна, where two consonants follow. In the genitive plural the [i] is definitely shorter than in the nominative. It seems, then, that the length alternation in vowels also prevails in parts of compounds which have a secondary stress which is derived from a full stress by the stress readjustment rule.

The phenomena described above can be accounted for along the following lines.

Let us postulate the following rules:

(14) A. A stress rule assigning primary stress

to the first syllable of every simple word and a secondary stress to every second syllable from there.

- B. A rule forming compounds by reducing the strength of the boundary between two or more words to be combined.
- C. A compound stress rule strengthening the stress of the first constituent of a compound and reducing the strength of other stress peaks. (This is possibly rule A. applied recursively.)
- D. The syllabification principle (12) applying to all stressed syllables, wherever it can.
- E. A length rule, applying to stressed syllables wherever it can, the domain being strings defined by the principle (12).

This mechanism would give the following derivation for the forms þjóðvís and leikvöllur. (There are a number of more or less unjustified assumptions made here about the underlying representations of the forms in question, but they are irrelevant to the problems discussed here.):

Underlying forms:

##þjoud##visa## ##leik##væ dYr##

The stress rule A. is applied to these forms, giving:

 1 1 1 1
##þjoud##visa## ##leik##væ dYr##

The length rule is applicable to both components of both forms, lengthening the vowel in the appropriate environments:

$\overset{1}{\#}pju:ð\overset{1}{\#}vi:ssa\#$ $\overset{1}{\#}lei:k\overset{1}{\#}vœdlYr\#$

(I assume that the syllabification principle applies automatically before the length rule to define its environments.)

These forms are then made subject to the compound-forming rule. This would give:

$\overset{1}{\#}pju:ð\overset{1}{\#}vi:ssa\#$ $\overset{1}{\#}lei:k\overset{1}{\#}vœdlYr\#$

The compound stress rule then applies, giving:

$\overset{1}{\#}pju:ð\overset{2}{\#}vi:ssa\#$ $\overset{1}{\#}lei:k\overset{2}{\#}vœdlYr\#$

Since we now have a new stress pattern, we can make the syllabification apply once more, pushing the syllable boundary as far back as possible from the vowel of the syllable bearing the heaviest stress. We can represent the output like this, disregarding the weakened word boundary, which may still be there:

$\overset{1}{\#}pju:ðv\overset{2}{\#}i:ssa\#$ $\overset{1}{\#}lei:k\overset{2}{\#}vœdlYr\#$

But the form for þjóðvísa does not now follow the rules of length distribution, so the length rule is applied once more. In the case of leikvöllur we can make it apply vacuously, since the form already has the right distribution of vowel length. The output will then be:

$\overset{1}{/}pju:ðv\overset{2}{i:ssa/}$ $\overset{1}{/}lei:k\overset{2}{vœdlYr/}$

Here again, there is hardly any need to stress the fact that this is far from being a permanent solution to the problems, but it does seem to me a plausible way of

approaching them. If such a framework is adopted, it becomes necessary to define the length rule more precisely than has been done so far.

In the derivation outlined above, the length rule functions as a sort of an 'everywhere rule' which is reapplied, whenever its structural analysis is met. It furthermore both lengthens and shortens vowels according to the environment. This can probably be represented most clearly by formulating the length rule as a two-sided transformation. It has been suggested by Lass (1974:322-323, see also references) that historical changes may be represented as a kind of two-sided rule, stating both what does happen and what does not happen at the same time. The rules can be said to have both a 'positive' and a 'negative' part. What I am suggesting is something similar, except that in this case both parts can be said to be 'positive' in that they imply changes, but these changes are in opposite directions and complementary, so to speak, one making vowels long and the other making them short. The length rule would then be stated in two parts, one part says that a vowel will be long if it is followed by one ^(or no) consonant within the same stressed syllable, and the other part says that a vowel will be short if it is followed by two or more consonants within the same stressed syllable. We can represent this rule as follows:

- (6) A. SD: $\bar{V} \quad \bar{C}^1 \quad \bar{C}^2 \quad \bar{C}^3 \quad \bar{C}^4 \quad \bar{C}^5 \quad \bar{C}^6 \quad \bar{C}^7 \quad \bar{C}^8 \quad \bar{C}^9 \quad \bar{C}^{10}$
SC: $1 \longrightarrow [+long]$

B. SD: V C₂ \$
1 2 3

SC: 1 \implies [-long] Condition: the syllable is stressed.

Part A. applies to all stressed vowels that are not marked [+long] before a single or no cosyllabic consonant, and part B. applies whenever a vowel is met that is not short before two or more cosyllabic consonants. As the rule is used in the derivation above, it applies both to vowels which have been marked with respect to the feature [+long] and vowels which are unmarked. This means that the rule in some instances adds a feature which is not in the input and in other cases it changes the value of a feature that already is in the input. This gives the length rule a character of an output condition, that is, whatever the input, the output conforms to the rules for distribution of length.

The attempt made above to describe the relationship of stress and length in compound words cannot be considered a strong argument for the assumption that stress is assigned to syllables before the length rule operates. What has been said above only supports that assumption in so far as it proves in the end to be the best way to account for the phenomena. The only thing we can say at this stage is that a model, which assumes that stress is assigned before the length rule operates, seems worth considering.

There is another argument that can be put forward to support the claim that stress precedes the length

rule. This has to do with what I will call contrastive stress. We have, so far, talked about stress as if it were a simple matter to state what it is. Not only is it difficult to find out what the phonetic correlates of stress are (cf. Lehiste 1970:106-142), but it is also a very complicated matter to determine what its linguistic function is in many cases and how to incorporate it into the description of the phonological or grammatical systems of languages. A distinction is often made between on the one hand what we may call a normal stress pattern, which is used when a word or a sentence is uttered in its most normal form without any emotive or stylistic overtones, and on the other hand, special uses of stress to emphasize or distinguish some parts of the utterance from other parts of it. What we have been discussing above can be described as the normal pattern of word stress in Icelandic. But use is frequently made of what Benediktsson (1963:148) calls 'morphological stress'. This is what has by some others (cf. Lehiste 1970:150-51 and references) been called contrastive or emphatic stress. We are dealing with such examples as: Bókin er á borðinu, (ekki á stólnum). 'The book is on the table, (not the chair)'. Here the word borðinu bears heavier stress than other words in the utterance, and this is done in order to contrast it with the other alternative, namely that the book is on the chair. This is a very clear example of the use of contrastive stress as I understand it. But, as Benediktsson

points out (loc. cit.), contrastive stress can also be used in Icelandic to contrast different inflectional forms of the same words as in Ég sagði 'gestinum', ekki 'gestunum' 'I said '(to) the guest, (sg.)' not '(to) the guests (pl.)'. Here, contrastive stress is put on the endings -inum and -unum to emphasize the difference between them. The interesting thing is that when the endings are stressed in this way, the vowels automatically lengthen, so that gestinum is something like [²ɛstI:nYm] and gestunum something like [²ɛstY:nYm], where the numbers once again represent relative strength of stress. Surely, it is the stress (whatever it is) that is used to contrast the endings, and bring them out of their context, but not the length of the vowel; and in that case the stress must be the conditioning factor for the length. By the same token, when endings, which have two consonants following the vowel bear such a contrastive stress, the vowel is not lengthened. This, I think, shows without any doubt that the length rule must be defined in terms of stress and stress be distributed before the length rule applies.

5. Some problematic forms

One fact concerning syllabification and length is still to be mentioned. According to the principle above, v, i, r following p, t, k, s are assigned to a following syllable: skrökva, setja, pukra 'to be secretive'. This presupposes that there is always a

following syllable to which the v, r or j can be assigned. This is indeed generally the case; sequences like tv, tj, tr etc. don't normally occur word-finally. There is, however, a set of exceptions to this. These are a limited number of nouns, generally derived from intransitive verbs of action. For example, corresponding to the above mentioned verb pukra, there exists a deverbative noun pukr 'secrecy, the act of being secretive'. Similarly there are pairs like sötra 'to sip' - sötr 'the act of sipping', kiökra 'to wail' - kiökr 'the act of wailing', sifra 'to lament', sifr 'the act of lamenting'. In these forms where we have a word-final tr or kr, there is no following vowel to connect the r with, so one would expect a syllabification like pukrs, and thus a short vowel according to the length rule. This is not the case, however; pukr, sötr, and kiökr all have long vowels.

It may seem that this is serious counterevidence to the analysis suggested above. I am not sure that it is, however. It seems that the forms in question are marginal in the language and their status in the system very special, and it can even be said of some of them that their wellformedness is doubtful. I am, for example, not at all sure that I can accept a form like lötr from lötra 'walk slowly'. In a way, these forms have a similar status to derived forms in English like the verb to comrade, in sentences like 'don't you dare comrade me!' (= 'don't you dare call me comrade'). The derivational

relationship is purely from one surface form to another, that is, the noun pukr is derived from the verb pukva, just as the verb to comrade is derived from the (speech act of uttering) the noun comrade. Of course, it can be said that this does not help the phonology, since the peculiarity just mentioned is morphological, and why should that affect the phonology? But the fact that these forms are morphologically (derivationally) special may act as an 'excuse' for them to go contrary to otherwise valid phonological generalizations.

Similar phenomena are mentioned by Kahn (1976: 121-124) from English. There are two generalizations that can be made about the distribution of low vowels preceding /r/ and nasals in American English:

- (a) Instead of [æ], orthographic a appears as [ɑ] in front of a cosyllabic /r/: car, (with [ɑ]) but carriage (with [æ]). (In the latter form, the /r/ begins the second syllable.)
- (b) [æ], orthographic a, is raised to something similar to [ɛ] in front of a cosyllabic n or m: can with [ɛ] (or something similar) vs. canon with [æ].

These generalizations could conceivably be set up as phonological rules for the dialects in question. But Kahn points out that in forms like Lar' derived from Larry and Jan', derived from Janice, these rules don't apply. Lar' has [æ] instead of [ɑ] and Jan' has [æ] instead of the raised variant. These phenomena seem to

be of the same sort as those we have seen from Icelandic: secondary derivational processes are allowed to lead to breaches of otherwise valid phonological generalizations. I will therefore conclude, tentatively, that these facts are not to be taken as direct evidence against the rules they break, but that they have to be dealt with in some special way. It may perhaps be said that they show that phenomena like the length rule in Icelandic and the 'syllabification principle', and the stress assignment rule connected with it, as well as the rules governing the distribution of [æ], [ɑ] and [ɛ] in American English operate at a relatively abstract level in the phonology, since they are not absolutely exceptionless. But then: are there any phonological regularities absolutely without exceptions?

Concerning the Icelandic examples, it can be added that the length rule (and the processes related to it) is not the only rule broken by forms of this sort. The forms pukr, sötr, kiökr and sifr (the last of which does not break the length rule) break another general rule, namely that $\underline{Cr}\{\# \atop C\}$ (C = any Icelandic consonant) does not occur in Icelandic. It has been proposed (Orešnik 1972) that there is active in Icelandic an epenthesis rule which inserts an /Y/ (orthographic y) in the appropriate environments. This would account for my (and many other people's) tendency to pronounce the forms in question with an 'epenthetic' /Y/: pukur [p^hY:k^hYr], kiökur [cœk^hYr] etc. Still another rule is broken by forms of

the same type. From the verbs grenja 'to cry, to howl' and hneggja 'to neigh', nouns like greni 'crying' [grɛnʝ] and hneggi 'neighing' [nɛc:] can be derived. The first of these forms breaks the rule that C_j is generally not allowed word-finally in Icelandic, and depending on the way the palatal in hneggja [nɛc:a] is analysed, the form either breaks the same principle or one forbidding a palatal in word-final position.

It is worth noting that all these rules broken by the deverbative nouns look very much like syllable structure constraints, and this may be the character of the length rule as well. Indeed, my last formulation of the rule (pp. 81-2) suggests this in a way, since it is basically an output condition, a well-formedness constraint on phonological forms. But it seems that this constraint may not hold on the most concrete of phonological levels.

Chapter III.
LENGTH IN OTHER GERMANIC LANGUAGES

1. Faroese

Faroese can be said to be the closest to Icelandic of the Scandinavian sister tongues. Many parts of the morphology and syntax are similar although, of course, there are notable differences. In the phonology, which is our concern here, there are also similarities, although here again the differences are substantial. Assuming that Icelandic and Faroese derive from a common variety of Nordic, it can be said that Faroese has shown a still greater tendency to diphthongize long vowels than Icelandic has. It is also notable that there is greater phonological dialect variation here than in Icelandic.

Modern Faroese phonology shows a (for us) important similarity to Icelandic in that vowel 'length' is predictable in stressed syllables on the basis of the following consonantism. (I will explain the quotation marks around the word 'length' in a moment.) The main rule is the same as in Icelandic, namely that vowels are short when followed by two or more consonants, but long otherwise. This indicates that Faroese has undergone a quantity shift like the Icelandic one.

As in Icelandic, there are exceptions to this rule of length distribution in that in most dialects there

are sequences of two postvocalic consonants that have long vowels preceding them. These are p, t, k, s, + i, r and p, k, + l (Zachariassen 1968:46 and Lockwood 1955:8). An interesting difference, compared to Icelandic, is the fact that pl and kl are preceded by long vowels (and have no preaspiration) whereas in Icelandic these are preceded by short vowels (and have preaspiration). It is also interesting that tl behaves differently from pl and kl in Faroese. It is probably no coincidence that kl and pl are permissible word-initial clusters whereas tl is not. It will make an interesting study to attempt an analysis of the length rule in Faroese in terms of syllabification and compare the results with Icelandic.

The above mentioned exception to the length rule is, however, not valid for all Faroese dialects. In the dialect spoken on the southernmost island of Suðuroy, vowels are short in front of all sequences of two or more consonants, including those that are exceptional in the other dialects (Zachariassen 1968:47). Thus, forms like vitja 'to visit' and vetrar 'winter' (nom.pl.) have short stressed vowels in the dialect of Suðuroy. There is, in this dialect as well as the others, a difference between the p, t, k, s + i, r and p, k + l sequences and other postvocalic sequences like tl, tn and kn in that preaspiration appears on the stop only in the latter, i.e. vetrar has a pronunciation something like [vɛtrar], but vatn 'water' something like [vəhtn]. This fact,

Zachariassen suggests, could perhaps be taken as indication that the shortness of the vowels in front of tj, kr etc. in the Suðuroy dialect is of rather recent origin, since it may seem that preaspiration arose historically on the stops p, t, k when preceded by short vowels. This, as well as the other particulars concerning the length rule in Faroese, provide interesting material for study, for which there is no room here. Anyway, it can be said that roughly the same situation prevails in Faroese as in Icelandic as far as the distribution of quantity is concerned.

To give a simple and reliable picture of the history of the Faroese vocalism is difficult, partly because of the lack of evidence and partly because the development seems to have been so complicated. To make things still more complicated, there are considerable dialectal differences, and I know of no comprehensive study of Faroese dialects. (See, though, Jakob Jakobsen's overview in Hammershaimb 1891:LVII-LIX.). Attempts at synchronic analyses are to be found e.g. in Bjerrum (1964), O'Neil (1964), S. Anderson (1972b), Taylor (1973) and Árnason (1976), and phonetic studies are to be found in Rischel (1964) and Hammershaimb (1891:LVII-LXIV). What I have to say about Modern Faroese is largely based on data from Tórshavn-speech, cf. Árnason (1976) and Lockwood (1955).

The most striking feature of Faroese, compared to Icelandic, is that the difference between 'long' and 'short' vowels as distributed by the length rule, is

much more qualitative here than in Icelandic. In particular there are five vowels, the reflexes of Old West-Scandinavian /u:/, /au/, /o:/, /ɑ:/ and /ǫ:/, /e:/ and /æ:/ (cf. below Chapter IV, Section 1.1), which show alternations between diphthongs in the long environments and monophthongs in the short environments. There are morphophonemic alternations between [au:] and [Y], (historical /u:/), [ɛi:] and [ɛ] (historical /au/), [ɔu:] and [œ] (historical /o:/), [ɛa:] and [a] (historical /a/, /e:/ and /æ:/), and [ɔɑ:] and [ɔ] (historical /ɑ:/). It seems that these alternations reflect three historical changes. (For work on the history of the Faroese vowel system, see e.g., Amundsen 1964, Chapman 1962:131-134 and Rischel 1968). Firstly, this shows that the old long monophthongs /u:/, /o:/ and /ɑ:/ have become diphthongs (in long environments at least). Secondly, it shows that the diphthong /au/ (Modern Icelandic [œy]) has become front and unround and lost its second component when short. Thirdly, it shows that the old short /a/ has become a diphthong in long environments. I suggest that these changes reflect, along with the quantity shift, three very basic processes that have affected Faroese stressed vowels: (1) A widespread diphthongization of old long monophthongs, (2) A monophthongization (loss of the second component) of diphthongs in the short environment of the length rule created by the quantity shift, and (3) a diphthongization of old short /a/ in the long environments of the length rule. In

addition to these, what I would call the major changes in Faroese vocalism, a number of mergers occurred, for example old short /e/ merged with /o/ in front of nasals, and /ø/ elsewhere, and old long /e:/ merged with /æ:/, which in turn merged with old /a/, giving [ɛa:] / [a] in the northern and central dialects, but [ɛ:] / [ɛ] in the southern dialect area. Also, there are some qualitative changes (apart from the ones already mentioned) which don't concern us here. To give a rough idea of the changes involved, I present here a correspondence table between (reconstructed) Old Faroese (basically the same as Old Icelandic) and Modern Faroese vowels (cf. Rischel 1968:109 and Árnason 1976:59):

Old	Modern		Dialect		
	Long	Short			
/i:/	ɔy:	ɔy			
/y:/					
/i/	i:	ɪ			
/y/					
/e:/	[ɛa:	a	Central		
/æ:/				a:	North
/a/				ɛ:	South
/e/	ɛ:	ɛ			
/ø/	œ:	œ			
/ø/					
/ø:/					
/u/	u:	ɔ			
/u:/	au	Y			
/o:/	[ou:	œ	Central		
				øu:	North
				ou:	South

Old	Modern		Dialect
	Long	Short	
/o/	o:	ɔ	
/ɑ:/	[ɔɑ: ɒ:]	ɔ	North
/ɛ:/		ɒ	
/ei/	ai:	ai	
/ey/	ɔi:	ɔi	
/au/	ɛi:	ɛ	

As can be seen from the table, diphthongization has taken place in the following old long vowels: /i:/ and /y:/, /o:/, /ɑ:/ and /ɛ:/, and /u:/. That is, these vowels have diphthongal variants in long position in most dialects: [ɔy:], [ɔu:], [ɔɑ:] and [au:] respectively. The old /e:/ and /æ:/ have a diphthongal long reflex in the central dialect, which might mean that they have undergone the same type of process (e:, æ: → əi: → [ɛa], or something similar), but having merged with an old short vowel (/a/), their status is somewhat special. (I will come to this later.) There are two possible ways of accounting for the short reflexes of these vowels. One is to assume that /u:/, /o:/ and /ɑ:/ ~ /ɛ:/ (and perhaps /æ:/) did not diphthongize in the shortening environments, but there were direct changes ai, ei → [ɔ] (or whatever), oi → [oe] and ui → [Y] in the appropriate surroundings. Another possibility is that these vowels diphthongized both in the shortening and lengthening environments, but ou, oa and au were later monophthongized in the shortening environments. In that case, it would be most natural to

assume that the long vowels in question were diphthongized before the shortening part of the quantity shift (the one making vowels short in front of two or more consonants, cf. Chapter IV below) became operative. In that case, the first vowel of húsfólk [hYs:fœlk] 'people of the house', which derives from old /u:/, will have developed something like this: ui → au by diphthongization, and then by shortening/monophthongization au → [Y]. Similarly the long /a:/ would have developed along the following lines in the shortening environments: ai — æ — [ɛ]. For the old /i:/ and /y:/, however, the shortening did not lead to a monophthongization, since the modern short reflex is diphthongal ([ɔy]). In this case there is a further complication in that there is a merger of an originally rounded and an unrounded vowel. It has been suggested that the roundness of the Modern Faroese diphthong stems from the /y:/, that is, that the result of the merger of /i:/ and /y:/ was a rounded vowel, which later diphthongized. (cf. Amundsen 1964:57-8 and Rischel 1968: 101-102). It is, incidentally, worth noting that, whatever else, the merger and common diphthongization seem to have preceded the quantity shift, since otherwise it would be difficult to explain the fact that both show the same (diphthongal) quality in shortening and lengthening environments. If the merger of /i:/ and /y:/ can be dated on the grounds of manuscript spellings, we thus have a possible terminus post quem for the quantity changes.

The hypothesis that the old long vowels that show diphthong-monophthong alternation first diphthongized in all environments and were later monophthongized in shortening environments derives some plausibility from the fact that the original diphthong /au/, which has the modern long reflex [ɛi:] has lost its second component in the shortening environments. Thus, it is necessary anyway to assume that a (post-quantity shift) monophthongization took place, and then it would be natural to assume that it affected the new diphthongs au, ou and oa (and perhaps ea) as well as ei. The exceptionality of [ɔy:] / [ɔy] in having a diphthongal short allophone would be explained in a way similar to [ai:] / [ai] (old /ei/) and [ɔi:] / [ɔi] (old /ey/).

The case of old short /a/, which along with old /e:/ and /æ:/ shows a diphthongal long reflex, is special. Here we have an old short vowel that has diphthongized and merged with old long vowels. This is unusual within the Faroese system, both since the general tendency is to keep old long and short vowels apart and since the other old short vowels have remained basically monophthongs (although some movement can often be detected in the long variants). A conceivable background for this situation is that old /e:/ and /æ:/, having merged, diphthongized to something like a which came later to appear as a monophthong in the short environments and,

after the quantity shift, old /a/ diphthongized when long and merged with the reflex of old /e:/ and /æ:/. This is very hypothetical and will need further justification before being accepted as a valid explanation. One fact that may be taken as indication that the modern pair [ɛa:] / [a] has a special historical background, is that it seems that (in certain dialects at least) it is the second part (more) of the diphthong that 'remains' in the short environment, whereas in the others, it seems to be the first component that remains ([au:] / [ʏ], [ɔa:] / [ɔ] , [ɛi:] / [ɛ]).

Rischel (1968:96) suggests that the 'quantity shift' took place in two steps in Faroese, by (1) a lengthening of short vowels before single consonants and (2) a shortening of long vowels in front of two or more consonants and that the shortening took place somewhat later than the lengthening. He does not, however, present any positive arguments for this relative chronology of the quantity changes. As will be suggested below for Icelandic (and Norwegian and Swedish) it seems natural to assume that the lengthening and shortening did not take place simultaneously, but I know of no evidence that can be put forth in support of one or the other of the relative chronologies. The fact that the short alternants of many diphthongal vowels appear as monophthongs cannot be used as an argument in this case, e.g. maintaining that the shortening manifests itself as a monophthongization in some cases, and is therefore likely to have

occurred later than the lengthening of old short vowels. There is no reason, even if the 'shortening' was really a monophthongization (loss of a final mora), to assume that it took place later than the lengthening and, anyway, there is always the possibility that the monophthongization was secondary, as is suggested above.

It seems to me, then, to be an open question what was the relative or absolute timing of the Faroese quantity changes, but it is clear that what happened must have been very similar to what took place in Icelandic.

2. Norwegian

As is well known, Modern Norwegian shows a great deal of dialect variation with respect to phonology. The dialect differences show that the phonological development from Old Norwegian, which must have been very close to Old Icelandic, has varied greatly. Quantity seems to have been no exception; as opposed to a uniform situation in Modern Icelandic, there is considerable difference between Norwegian dialects as far as the quantity situation is concerned. The situation in the modern dialects can give us extremely valuable indirect evidence of how quantity developed in Norwegian, and it is therefore useful to make the present state of affairs our starting point.

In most Norwegian dialects the distribution of length is basically the same as in Icelandic, Faroese

and most Swedish dialects. We have long vowels occurring in stressed syllables in front of single consonants, in front of hiatus and in word-final position. Short vowels occur under stress in front of two or more consonants and also in front of geminates (long consonants). The types of stressed syllables that occur are then: V:C, V:ʃ, V:# and VC₂, where C₂ stands for two or more consonants or a long (geminate) consonant. This general situation shows that a quantity shift has taken place, given that stressed syllables of the type VC, and V:C₂ occurred in Old Norwegian, as in Old Icelandic.

If we start by looking at the Old Norse short syllables of the type VC, we see that the Modern Norwegian dialects show differing reflexes of these. Some dialects have (in some cases) eliminated this syllable type by lengthening the consonant, whereas in other cases the vowel has been lengthened. As we have seen, Faroese (and Icelandic) eliminated this syllable type by lengthening the vowel (there seem to be some exceptions to this in Faroese, cf. e.g. summar Common Nordic sumar 'summer') so already in this respect Norwegian distinguishes itself from the other West-Scandinavian dialects. The general rule for Norwegian is that the northern dialects and to some extent the eastern ones show a tendency to lengthen the consonant, whereas the southern and western dialects favour a lengthening of the vowel (see Indrebø 1951:221). Thus we have for example in the dialect of Trøndelag (near the town Trondheim) vatte

(ON vita 'know') and in Tromsø in the far North vette, as opposed to vi·ta in Sogn (near Bergen). (These forms are taken from Christiansen 1946-48:130.) The distribution of consonant lengthening vs. vowel lengthening corresponds well geographically with the distribution in Swedish, where consonant lengthening is mainly a characteristic of the northern dialects (Sveamål), but vowel lengthening a characteristic of the southern dialects (Götamål). This alone shows that the quantity shift was not as uniform in Norwegian (and Swedish) as it seems to have been in Icelandic and Faroese, judging from the situation in the modern languages. The Modern Norwegian situation shows that at least two changes could affect the ON short syllables of the type VC, one lengthening the vowel and the other lengthening the consonant. These changes are obviously mutually exclusive, that is, where one occurs the other can not, but it can be said that they aim at the same results, since they both lengthen previously short syllables. There are dialects which lengthen vowels in some environments but consonants in others (cf. Christiansen 1946-48:132), but one can generally say that consonant lengthening is a northern (and eastern) feature and vowel lengthening a southern one.

Even though the general situation in Modern Norwegian is that the Old Norse short syllables have been lengthened, there are exceptions to this. In the dia-

lect of Tinn in Telemark (in the South, west of Oslofjorden) bisyllabic forms like viku 'week', vyta, 'know' and smaka 'taste' with a short first vowel and even a 'quantitative and expiratory overweight on the second syllable'¹⁾

(kvantitativ og ekspiratorisk overvekt på etterstavingen) are reported to occur (Christiansen 1946-48:132). On the other hand the monosyllables seem to have lengthened in this dialect. The fact that a lengthening in monosyllables and a lengthening in polysyllabics don't necessarily cooccur shows us that the lengthening of stressed vowels that has hit Norwegian can be split up into two changes, which we can state informally like this:

(1) $V \longrightarrow V: / -C$

(2) $V \longrightarrow V: / -CV$

In most dialects which lengthen vowels, both of these changes have been completed, but in the Tinn dialect, only the first one has occurred. The Tinn dialect seems then to be a conservative variety of Norwegian with respect to the development of quantity. But there is a still more conservative dialect. This is the one spoken in northern Gudbrandsdalen. Here, both monosyllables and bisyllables retain ON short vowels in stressed syllables: lās (ON las) 'read' (past tense) lās (idem pres. tense) (Indrebø 1951:221). In this dialect, neither of the two vowel lengthenings have taken place.

A phenomenon worth mentioning in connection with the development of the ON short syllables is the so-called 'vowel balance'. This is a feature that is often used as an isogloss distinguishing between the two major dialect

areas in Norway. (The two main dialect areas are the West on one hand and the East on the other. The Eastern dialects show reflexes of the vowel balance, whereas the western dialects don't.) A distinction is made between ON bisyllabic words which were 'balanced' and those which were 'overbalanced'. The balanced words were the ones with a short first syllable, like for example vita 'know' and dagar 'days'. Here a 'balance' is said to have prevailed between the two syllables, since their length was similar. The overbalanced words were those with a long first syllable: høyra (OI heyra) 'hear', kasta 'throw' and blása 'blow'. In the vowel balance areas the vowels of the second syllables developed differently according to whether the first syllable was short or long, i.e. whether they were balanced or not. The result varies according to dialects, but in all vowel balance areas the second vowels of balanced words showed more resistance to weakening or deletion than in the overbalanced words. For example in the southern part of the eastern region we get ve.ta (ON vita), with a final -a retained as opposed to kasta (ON kasta) with a final -a 'weakened' to -e. Similarly, in Trøndelag, further north in the eastern region, we have vatta (ON vita) with a retained vowel as opposed to kast (ON kasta) with apocope. Western (and also northern) dialects on the other hand show the same treatment of the second vowel irrespective of the historical length of the first syllable. This we can see, for example, in the Sogn

dialect. Here we have vi·ta and kasta both with a final a retained; and in the dialect of Salta in northern Norway we get vet, kast with apocope in both forms. (The data are again taken from Christiansen 1946-48:130-131.) Christiansen (op.cit.:119) considers the retention (or resistance to weakening) of the second vowel in the balanced words to be caused by the fact that both syllables of the balanced words carried equal stress, or weight. As support for this hypothesis Christiansen cites the above mentioned forms vikú', vytá' and smaká' with a heavier stress and a longer vowel in the second syllable. The argument is presumably that it is easier to explain the fact that stress is on the second syllable, if it previously was not inferior to the first syllable as far as stress or 'weight' is concerned.

It is hard to say just what effect the vowel balance phenomenon can have had on the development of the quantity in short syllables, but it may not be a coincidence that both Tinn and the northern part of Gudbrandsdal are in the vowel-balance areas. If it is true that the balanced words had a different stress pattern from overbalanced words in some dialects, it is conceivable that this had some influence on the development of quantity. It is, for example, possible that the conservatism of the Tinn-dialect, where bisyllabic balanced words still have short vowels, can be explained in this way. If the first vowel did not bear heavier stress than the second one, there may have been no conditioning factor for lengthening it.

It is more difficult to explain the situation in northern Gudbrandsdal - where monosyllables also have short vowels - in these terms, since monosyllables like las had only one syllable to place the stress on, if there was any stress to be placed at all. Anyway, the vowel balance did not preclude lengthening of a short first syllable, since the lengthening has taken place in all vowel balance dialects except the ones I have mentioned. I will not draw any conclusions about the relation of vowel balance to the development of length from the data described above, since I have not had an opportunity to study these matters in detail, and the description given above is inevitably an oversimplification of the facts.

Another phenomenon that may bear some relation to the whole matter of quantity is the word tones. These word tones, usually called Accent 1 and Accent 2, are different prosodic contours of words, mainly based on pitch variation (cf. e.g. Gårding 1973:30-46). Different words have different tones or accents in most Swedish and Norwegian dialects, and minimal pairs have been cited to show that they are distinctive even though some scholars don't think they are (cf. Haugen 1967). The historical origin of these tones is probably that Common Nordic bisyllabic (and polysyllabic) words had different pitch (and perhaps stress) contours from the monosyllables. At the oldest stage this difference in contours was probably predictable from the number of

syllables in the words; polysyllables had the contour that later became Accent 2, but monosyllables had the contour that was to become Accent 1. (Cf. Oftedal 1952: 219 and 221-222). Later, when some monosyllables became bisyllabics by, for example, the affixation of the definite article (dag+inn > daginn 'the day') and the development of epenthetic vowels (before final liquids or nasals (ON akr > aker 'a cornfield'), these new bisyllabics still retained the same Accent 1 contour. Now some bisyllabics had Accent 2 and others (the new ones) had Accent 1, and the distribution of accents was no longer predictable from the number of syllables.

It is interesting to see whether there can have been some connection between the development of quantity and the accents. The data from the Tinn dialect, as mentioned above, seems to indicate that the lengthening of short monosyllables preceded the lengthening in polysyllabics, and, as was the case with vowel balance, it is conceivable that Accent 2 had something to do with this. The majority of bisyllabic words had Accent 2 at the time when the lengthening started taking place, and the conservatism of bisyllabics could then perhaps be ascribed to the fact that they had Accent 2. This could be made more plausible by observing that it is a general characteristic of the modern Accent 2 that it has a relatively late pitch peak, which could give the second vowel of a bisyllabic form more prominence than it would otherwise have. (Cf. Gårding 1973:44). It is also

possible that the vowel balance and Accent 2 were inter-related and that they both combined to make polysyllabics resistant to the quantity shift.

There is one type of words which could help to decide whether Accent 2 had any effect on the development of quantity, namely the bisyllabic forms with Accent 1. If Accent 2 tended to prevent first syllable lengthening, the Accent 1 words should have followed the monosyllables. A thorough investigation of the dialect material is needed in order to decide this, and I have found no allusion to this in any of the reference books I have seen on Norwegian. There is, however, some evidence to the contrary in that in some Swedish dialects in which the quantity shift has not been completed, the Accent 2 bisyllabics show a greater tendency to lengthen the first vowel than the ones with Accent 1. Söderström (1972: 91-92) cites examples from Luleå dialects in Sweden which show this. There are pairs like beka, v. with Accent 2 (grave) and a long vowel vs. beke (definite sg.) with Accent 1 (acute) and a short vowel. This seems, if anything, to indicate that Accent 2 makes a favorable rather than unfavorable environment for the lengthening of the first vowel. We will return to this briefly in section 3.

If we turn now to the Old Norse 'overlong' syllables, i.e. the type V:C₂, we see that this type has generally been excluded in Modern Norwegian. This shows that a historical change something like

$$(3) \quad V: \text{ --- } V / \text{---} C_2$$

has taken place. But here again, there are exceptions. In the dialect of Setesdal (in the south) forms with a long vowel or a diphthong followed by a long consonant are reported to occur (Indrebø 1951:222): no^utt, lⁱtt^a (ON nótt, 'night', léttr, 'light'). This shows that the change (3), just as (1) and (2), has not yet been completed in all dialects.

To sum up, then, we see that in Norwegian the following four historical rules affecting stressed (or first) syllables have operated:

- (4) a. $V \longrightarrow V: / \text{--- } C\#$
- b. $V \longrightarrow V: / \text{--- } CV$
- c. $C \longrightarrow CC / \check{V} \text{---}$ (Consonant lengthening)
- d. $V: \longrightarrow V / \text{--- } C_2$

From the sketchy picture presented above we see that the 'quantity shift' in Norwegian cannot have been a single, sudden turnover (omveltning), but rather a set of changes, which hit different dialects at different times, and in different ways. Sometimes consonants are lengthened and sometimes vowels, and some dialects have to some extent retained the old prosodic structure. In view of this, one must ask whether the term quantity shift is appropriate. Why would we want to group these rules together under a common term? The reason is, of course, that the overall effect of these changes is to change the rhythmic structure of the language so that all stressed (first) syllables are of the same length, i.e. either a short vowel + two or more consonants

(assuming that long consonants can be analysed phonologically as geminates), or a long vowel + no more than one consonant.

When two or more apparently separate rules behave in this way, that is, giving a unified and simply storable result, the term conspiracy has been used in synchronic phonology (cf. Kisseberth 1970). In a recent paper, Roger Lass (Lass 1974) has suggested that similar things appear in historical development. He sees in the development of quantity in English and Scots a gradual tendency to make vowel length predictable rather than phonemic. This tendency manifests itself in a number of apparently unrelated changes, which take place at different times in the history of the English dialects in question. Lass calls this 'linguistic orthogenesis'.

If the terms conspiracy or orthogenesis are to be applied in historical linguistics, the development of quantity in Norwegian seems to fit the terms extremely well. We have changes taking place at different times, which aim at a simply storable result. It would then seem to be proper to use these terms to denote the quantity changes in Norwegian, rather than using the term quantity shift, which seems to imply that a sudden revolution took place. But even though we adopt the terms conspiracy or orthogenesis, it does not necessarily mean that we have given a satisfactory account or an explanation of the facts. Inventing a name for things is, of course, not the same thing as stating what they

are. In the case of Norwegian (and the other Scandinavian languages) the quantity conspiracy can perhaps be explained in very down to earth terms in the following way. It is not an unnatural thing that stressed parts of utterances tend to become phonetically longer than their underlying structure may imply. In the case of the old short stressed syllables, this may have resulted in two, more or less accidentally distributed phonetic changes: a lengthening of the vowel or a lengthening of the consonant. Between generations these phonetic data get reinterpreted time and again, and the underlying grammars of younger generations may be slightly different from the grammars of older generations, until at some stage the (once perhaps irregular) phonetic alternations reach a firmer status in the language system. These systematizations may occur gradually. For example rule (4a) may become a part of the grammar of some dialect at an earlier stage than rule (4b). When the stage is reached, where both (4a) and (4b) are incorporated into the system, a language learner may make the generalization that stressed syllables are all long. This could be a very simplistic explanation for the disappearance of old short stressed syllables.

A different historical accident may have eliminated the old overlong syllables. The phonetic reason for this change may have been that long vowels tended to be shorter than predicted by their underlying forms when

followed by more than one consonant. A phonetic alternation like this may have been reinterpreted by younger generations, incompletely at first, until a generalization that phonemically long vowels are shortened before two or more consonants reaches the status of some kind of a phonological rule in the language. When these two, more or less accidental changes in the language, the lengthening of short vowels and the shortening of long ones, are completed, one can imagine a reinterpretation of the facts by a new generation of speakers who make the generalizations that all stressed syllables are of the same length and the length of vowels is predicted by the following consonantism.

3. Swedish

The quantity system of Modern Standard Swedish is the same as those of Standard Norwegian, Faroese and Icelandic as far as length in stressed syllables is concerned; that is, stressed vowels are long when followed by a single consonant or a vowel, and short when followed by two or more consonants (including geminates). As is the case with Icelandic, it has been a matter of dispute how to analyse this synchronic situation phonologically, that is, whether the phonemic length belongs to the consonants (Eliasson and La Felle 1973) or to the vowels (Elert 1964:12-46). We will not be directly concerned with that problem here, but will look briefly at the phenomena from the historical point of view.

Given that Swedish derives from a common Nordic ancestor with distinctive vowel length and stressed syllables of varying length, that is, short (VC), long (V:C or VCC/VC:) and overlong (V:CC/V:C:), we see that a quantity shift has taken place, since no short and no overlong stressed syllables are to be found in Standard Swedish. As we have seen, the quantity shift in Norwegian was not nearly as regular as the one in Icelandic seems to have been. Whereas the Icelandic quantity shift, generally speaking, only hit vowels, i.e. short vowels are lengthened and long ones shortened according to the environment, some Norwegian dialects sometimes lengthen consonants in old short syllables. The same is true of Swedish. In many northern dialects the consonant is often lengthened if the vowel is non-low and the consonant is /p,t,k, or s/. The more general rule for Swedish, however, is to lengthen the vowels. The different development of old short syllables as far as lengthening of vowels or consonants is concerned often shows up in Standard Swedish. Thus, in Standard Swedish we get gata [gɔ:t^ha] 'street' with a lengthened vowel (cf. Old Icelandic gata 'road') as opposed to vecka [vɛk:a] 'week' with a lengthened consonant (cf. OI vika 'week'). Geographically the main rule for Swedish, as for Norwegian, is that the southern dialects tend to lengthen the vowel, whereas the northern dialects have a tendency to lengthen the consonant according to the rules mentioned above (Wessén 1945:60-62).

Apart from this variation concerning the lengthening of consonants vs. the lengthening of vowels, there is in Swedish dialects a further irregularity with respect to the development of quantity in that, as in Norwegian dialects, the quantity shift has not everywhere reached its final stage.

Söderström (1972) describes Swedish dialects which have, to a varying extent, retained old short syllables. A striking feature concerning the retention of old short syllables is that a considerable difference shows up, according to whether the old short syllable is in a monosyllable or a disyllable. Monosyllables show a greater tendency to lengthen their only syllable than do the disyllables their first syllable (Söderström 1972:88). The areas that Söderström's study covers are the following: 1) Överkalix and Nederkalix in the far northeastern part of Sweden; 2) The area around Piteå, further south on the East coast (both of these are in the Norrbotten region); 3) Nordmaling, still further south on the coast, a bit south of Umeå in the northern part of Ångermanland; and 4) Ragunda in Jämtland. All of these dialects are within the larger area of Norrland. The first three of these dialects have still largely retained old short first syllables in bisyllabics, while mostly having lengthened monosyllables. This can give inflectional paradigms where there is a morphophonemic alternation in the same word between short vowels and long vowels or diphthongs

according to whether a syllable (an inflectional ending) follows or not. Thus, the nominative sg. of Standard Swedish väv 'a cloth' has in Överkalix, Nederkalix, Piteå and Nordmaling a long vowel or a diphthong /ve:iv/, /ve:v/, /vɛ:v/, whereas the plural, with the ending /-a/, has a short first syllable: /veva/, /veva/. (Söderström 1972: 129). The Ragunda dialect (4) seems to be not as conservative as the other dialects mentioned, since only bisyllabic words with Old Swedish /i/ and /u/ are reported from that dialect with short first syllables. It seems that the Old Swedish low vowel /a/ shows the greatest tendency to lengthen, whereas the high vowels show more resistance, for example the Old Swedish word bit 'bite' shows up in Överkalix as bed with a short vowel (Söderström 1972:58), whereas Standard Swedish mat (Old Icelandic matr) 'food' shows up as med with a long vowel (Söderström 1972: 57).

These synchronic facts give strong evidence to the effect that the lengthening of old short syllables took place first in monosyllabic forms and later in disyllables. This agrees well with the statement made by Noreen (1904:123) that signs of the lengthening in monosyllables had become general in Swedish manuscripts after 1350, whereas clear signs of the lengthening in bisyllables are not older than ca. 1500. This, furthermore, conforms well to what we have already said about Norwegian dialects, the Tinn dialect retaining short vowels only in polysyllables, but the apparently more conservative dialect of North Gudbrandsdalen retaining a short vowel both in monosyllables and disyllables.

One can hypothesize from this that it is a common feature of all Norwegian and Swedish dialects that they have lengthened short monosyllables before they lengthened the first syllables of bi- and polysyllables. If this is correct, the northern Scandinavian dialects, Norwegian and Swedish show different behaviour from Danish in the South, since, as we shall see (Section 4.), old short monosyllables were never lengthened in Danish, only the first syllables of polysyllables.

I do not claim to be able to explain here why these subsets of Scandinavian dialects, i.e. Danish on the one hand and Norwegian and Swedish on the other, developed differently in this respect, but only make a few suggestions. In Section 4. and 6. it is suggested that the consonant shortening (degemination) that occurs in Danish and disrupts the development of vowel quantity may be the same that hit German and English, and it seems not unlikely that a contact with Southern or Western Germanic people may be responsible for this other piece of peculiar behaviour on the part of Danish. To draw any conclusions about this, one must of course make a careful study of the chronology and geographical distribution of the phenomena involved.

If we (tentatively) ascribe the peculiar development in Danish to West/South-Germanic influence, a natural corollary of that would be to say that the genuinely 'Nordic' way of lengthening old short syllables is to start with monosyllables, i.e. having the chronological order:

1. $V \rightarrow V: / \text{---} (C)\#$
2. $V \rightarrow V: / \text{---} (C)V$

This seems to be supported by the facts in Swedish and Norwegian dialects. As to why the quantity shift took this form in Norwegian and Swedish, a number of relatively plausible explanations can be proposed, but they may turn out to be difficult to choose between, let alone prove.

As mentioned in connection with Norwegian, two Continental Scandinavian (as opposed to Danish) features look as though they may have had some relation to the development of quantity. These are the so-called vowel balance and the word tones. Related to both of these is probably stress, its placement and nature. A fourth phenomenon that may be (and probably is) related to this whole business is the so-called jamning (Norwegian) or tilljämning (Swedish). This is a vowel assimilation between the first and the second vowels of bisyllabic words. It can be both progressive and regressive, that is, we can either get, e.g. in Norrlandic Swedish, lovo from Old Nordic lofa 'to praise' with the second vowel assimilating to the first, or yuku from Old Nordic viku 'week' (accusative sg.) with the first vowel assimilating to the second (Bergman 1973:106). This assimilation only takes place in words with old short first syllables and is most prominent in northern and western Norrlandic dialects in Sweden. All of these phenomena can be said to indicate a certain balance between the first and the

second syllables of bisyllabic words with short first syllables.

Perhaps the least likely of these phenomena to be connected with the development of quantity is the tones. Firstly, there does not seem to be any difference in the tonality of bisyllabic words according to whether they have old long or short first syllables. This in itself, of course, does not prove that it could not have had some special effect on the old short syllable words, but there is no compelling reason to assume that it should have either. Secondly, the above mentioned data from the Nederluleå dialect (Section 2., cf. Söderström 1972: 91), namely the pair be`ke with Accent 2 and a long vowel vs. be'ke with Accent 1 and a short vowel seem to suggest, if anything, that length goes with Accent 2 (which is original on bisyllabic forms) and that shortness goes with Accent 1 (which is original on monosyllables). In view of this, it seems unlikely that Accent 2 by itself caused the bisyllabics to retain their original short syllables longer than the monosyllables. Thirdly, it may be mentioned as indicating that the length phenomenon and the word-tones are relatively unrelated, that there seems, as far as is known, to be no difference between the pitch variation in Accent 2 words with short first syllables and ones with long first syllables in dialects which have retained the length difference (See Gårding 1973:34 and references.)

The other phenomena, namely vowel balance, tilljämning and stress remain candidates for contributing to the explanation of why the bisyllabics retained short syllables longer than monosyllables. But in dealing with these, it must be borne in mind that they may all be aspects of the same phenomenon, and they are more than likely interrelated in some way, although it may be difficult to maintain that one is a consequence of the others or one is the cause of the others.

Both the vowel balance and the tilljämning are characteristic of northern dialects (Bergman 1973:104-105, and Wessén 1980:50-52), and we have seen that within the Swedish dialect area, it is in the northern dialects that we find retained old short syllables, with the bisyllabics more conservative. We mentioned (Section 2.) that it had been suggested that the vowel balance in Norwegian had been connected with a relative equality of stress on the first and the second syllable of 'balanced' bisyllabic words. One can perhaps say that tilljämning represents a similar tendency, that is, the stressed and the unstressed syllables assimilating to each other and there being no clear sign of one 'dominating' over the other. If both of these phenomena, the tilljämning and the vowel balance, are related to a relative balance of stress between the first and the second syllable in bisyllabic words with a short first syllable, one might wonder whether the relative lateness of lengthening of old first syllables in bisyllabics could be related to this general situation.

4. Danish

As mentioned above, Danish shows some unscandinavian features with respect to quantity. It can, for example, be maintained that vowel quantity is distinctive in Danish, whereas the other Scandinavian languages have lost that feature. In this section I would like to give a brief survey of the development of quantity in Danish by way of comparison with the other Scandinavian languages.

There can hardly be any doubt that the Danish quantity system goes back to the same one as the other Scandinavian languages. I would like to have a look at the history of the Danish system and try to find out just where the difference between Danish and her sister tongues lies. I will try to show that Danish underwent a part of the general Scandinavian quantity shift, but that its development was disturbed by a change of another kind, which shortened all long consonants and thereby destroyed an important part of the environment which in the other languages determines vowel quantity.

We can take as our starting point the situation in Common Nordic, where there are usually considered to be syllables of three lengths. (I pick the examples from Old Icelandic, but that should not make any difference):

(1)	short:	VC	<u>man</u> 'slave'
		V:	<u>má</u> 'am allowed to'
	long: [VC:	<u>mann</u> 'man' (acc.)
		VCC	<u>gest</u> 'guest' (acc.)

overlong:— [V:C: spánn 'spoon'
 V:CC ást 'love'

The situation in Modern Standard Danish (MD) can be described roughly as the following (cf. Rasmussen 1972:57):

(2) Monosyllables

VC hat [hæt^h] 'hat'
 VCC hest [hest^h] 'horse'
 V'C pén [p^hɛ'n] 'nice'
 VC' pen [p^hɛn'] 'pen'

The acute accent is here intended to represent the stød, which can fall either on the consonant (pén) or the vowel (pen) in monosyllables. (On the phonetic nature of the stød see e.g. Jespersen 1922:118-19 and P. Andersen 1954:320). The stød is probably a surface realisation of underlying length in the vowels, and perhaps in the consonants too. (Of course this only applies to the voiced consonants, since the unvoiced ones, e.g. /s/, 'cannot' take stød (cf. e.g. Jespersen 1922:156).) As support for the analysis of stød as a surface marker of underlying length in the vowels we can cite the fact that 'stød-less' dialects show long vowels, where the 'stød-dialects' have stød on vowels. In the South-Sjælland dialect the difference between pén and pen is in vowel length, the former having a long vowel, the latter a short one. Another fact, perhaps more important, is that there occur morpho-phonemic alternations between a stød-vowel and a long one: man [mæ'n] 'conjure' (imperative) mane [mæ:nə]

(infinitive). The reason for this is that the stød does not (generally) appear in bisyllabic words, cf. the following:

(3) Bisyllabics:

VCV	<u>falde</u>	[falə]	'fall'
VCCV	<u>hente</u>	[hɛnt ^h ə]	'fetch'
V:CV	<u>male</u>	[mæ:lə]	'paint'
V:CCV	<u>hoste</u>	[hɔ:stə]	'cough'

This looks very much like a system with phonemic vowel length, which is in some dialects realised in monosyllables as stød. The following minimal pairs fit that analysis very well:

- (4) hale [hæ:lə] 'tail' vs. halve [hælə] '(the) half'
hvile [vi:lə] 'rest' vs. vilde [vilə] '(the) wild'
hvil [vi'l] 'rest' (imp.) vs. vild [vil] 'wild'

But things are not all that simple, because there are examples which show an alternation between long and short vowels in the same morpheme, and in some environments a phonemic distinction between long and short vowels is impossible. Examples showing morphophonemic alternation between short and long vowels are:

- (a) tabe [t^hæ:bə] 'loose' tabte [t^hæptə] 'lost'
(This example represents the fact that long vowels don't occur in front of consonant clusters, except /st, sk, bl, bn/.)

- (b) mæð [mæð] 'food, feed (imp.)' mæde [mæ:ðe] 'to feed'
This second example reflects the fact that in the environment -CV short vowels sometimes become long (or take the stød). This is not a general rule, however, since we

have examples like vind [vin^ː] 'win' (imp.) and vinde [vine] 'to win' with no length alternation. It is perhaps possible to make the rule more general by analysing the postvocalic consonant in vinde as underlyingly long, but this is a problem in the synchronic phonology of Modern Danish, which I am not directly concerned with here. (See e.g. Hjelmslev 1951/1973 and Basbøll 1970-1971 on this matter.)

What these examples are meant to show is that in MD there seem to be two rules that can be traced back to quantity changes of the sort that hit the other Scandinavian languages:

- (5) (a) $V \rightarrow \check{V} / \text{--- CC}$
(b) $V \rightarrow V: / \text{--- CV}$

That is, these rules, which are generally valid for MD, can be taken to reflect a historical shortening in closed syllables and a lengthening in open syllables. And if we look into the history of Danish, we find exactly these changes taking place. Rasmussen (1972: 63) describes the two following quantity shift rules:

- (6) (a) $V \rightarrow V: / \text{--- CV (ca. 1300)}$
(b) $V: \rightarrow V / \text{--- CCV (15th century)}$

The MD alternations, whatever the synchronic rules may be that are needed to account for them, are quite clearly reflexes of these historical changes.

The fact remains that MD shows striking dissimilarities from the other Scandinavian languages in that vowel quantity/stød seems quite clearly to distinguish

between minimal pairs, whereas in the other Scandinavian languages this is generally not the case. Admittedly, arguments have been put forth for vowel quantity being phonemic in the other Scandinavian languages, but there are no phonetic minimal pairs, where the vowel quantity seems to be the only distinctive factor, since short vowels are ^{always} followed by clusters or long consonants in a stressed syllable, and therefore arguments can be held in favor of the vowel quantity being redundant. This is much more difficult, if not impossible in Danish.

The reason for this difference between MD and the other Modern Scandinavian languages is perhaps that Danish underwent a general shortening of long (geminated) consonants. The dating of this change seems to be disputed, mainly because there is little or no orthographical evidence for it. Skautrup (1944:254) dates it as early as 1300, but Rasmussen (1972:167) seems to date it later, as late as the beginning of the 16th century. It must have taken place later than the change (6a), since otherwise the form vilde should have a long vowel. Whatever the dates of these changes, the consonant shortening has neutralised a large part of the environment, which in the other Scandinavian languages determines the length of the preceding vowel. This can be illustrated by the word pair vild 'wild' (Old Icelandic villr) in the definite form vilde (OI villi, Old Danish villæ) and hvile 'rest' (OI hvila, OD hvilæ). The respective OD forms must have been approximately [vil:ə] and [(h)vi:lə].

These forms were unaffected by the two changes in (6), but then a change took place, which can be stated informally like this:

(7) C: \rightarrow C / V —

This change seems to have been general and affected all long consonants. After this we can hypothesise a situation, where the forms are (I use the MD orthographic forms as references):

vilde [vilə] hvile [(h)vi:lə]

For these to be kept apart phonologically, the natural thing to happen would be to phonemicise whatever difference there is between the vowels, and this is of course the difference in quantity. A similar thing, but slightly more complicated, happened to the monosyllables. We can take vild (the indefinite form of the same adjective) and hvil (imperative) as examples. The thing to happen with these forms according to the rule (7) would be [vil:] \rightarrow [vil] and [(h)vi:l] \rightarrow [vi:l], and this is actually the situation in the stød-less dialects of MD. But other dialects use the placement of the stød, which is considered to be the historical reflex of Accent 1 (which was restricted to monosyllables and is still existent in Norwegian and Swedish), to distinguish between these forms: vild [vilˀ], hvil [viˀl]. I am not sure how to explain this in a natural way, but it occurs to me that the stød-dialects have also undergone phonetic shortening of the vowels in monosyllables, in which case there is nothing

left to distinguish the two forms by except, perhaps, the peak of the old Accent 1, which may have been in a different place in words of different syllable structure, for example on the consonant in VC: types, but on the vowel in V:C types. This shortening of the vowel in monosyllables would be the historical origin of the synchronic phonological rule set up by e.g. Hjelmslev (1951/1973) and Basbøll (1970-1971), making stød a surface marker of underlying length.

If Danish has developed in the way I tried to indicate above, a quantity shift has started to hit Danish, perhaps in a slightly different way from the other Scandinavian dialects. But before the quantity shift could be brought to its natural end, Danish was hit by a change of an 'unscandinavian' type, which blocked the way for further development along the same lines. Why this happened in Danish in particular and not the other Scandinavian dialects, I am not absolutely sure, but it is perhaps not a coincidence that German and English have no long consonants either. It seems to me to be quite likely that there is a connection.

5. Gothic

It has been a matter of lengthy dispute, whether Wulfilian Gothic had distinctive vowel length or not. Some scholars of the structuralist school have argued that the Gothic vowel system did not have length as a distinctive feature. The main argument is that the synchronic evidence, mainly the graphemic system used

in extant manuscripts of Wulfila's Bible translation does not show directly that the length distinction prevailed. There is, for example, no distinction made in the spelling between PGmc. /u/ and /ū/ in forms like sunus 'son' and brups 'bride' which on comparative grounds can be reconstructed with historically short and long vowels respectively (cf. Old Icelandic sonr/sunnr vs. brúðr). There is, however, one historical long vs. short distinction which is consistently made in Gothic spelling, namely that between PGmc. /i/ and /ī/, spelled i and ei respectively: greipan 'catch' (infinitive) (OI grīpa) vs. gripum (idem past 1. pers. plural) (OI grīpum). This spelling difference has been interpreted by those who maintain that the length distinction was lost in Gothic as representing a difference in quality rather than quantity, i.e. that the length opposition had been replaced by a qualitative opposition in these vowels. As a representative of those who maintain that vowel length was non-phonemic in Gothic I would like to cite Marchand (1973). He sets up the following vowel system (p.95):

		Front	Central	Back
<u>High</u>	Close	i		u
	Open	ī		
<u>Mid</u>	Close	e		o
	Open	ɛ		
<u>Low</u>			a	

/i/ is represented in the orthography by ei, /ī/ by

i, /e/ by e, /ε/ by ai, /a/ by a, /u/ by u and /o/ by au. In addition to this, Gothic had the diphthong /iu/, written iu. As can be seen, Marchand assumes that the PGmc diphthongs /au/ and /ai/ had been monophthongized into /o/ and /e/ respectively.

If we assume that pre-Gothic Germanic had the following vowel system (cf. e.g. Prokosch 1939:99-105):

Short		Long		Diphthongs		
i	u	ī	ū	ai	au	iu
e		ē	ō			
a						

the Gothic situation claimed by Marchand presupposes the following changes:

- 1) Merger of i, e > i
- 2) Monophthongization: ai > ē, au > ō
- 3) Breaking: i > ε, u > o / —h, h, r
- 4) Quantity shift

The order in which these changes are listed above would probably not necessarily reflect their chronological order, but basically these should be the effects, when Gothic is compared with PGmc. The quantity shift is a (perhaps context free) loss of length as a distinctive feature in the vowels resulting in a merger of all long-short vowel pairs (some of which reflected partly PGmc monophthong-diphthong opposition, e.g. PGmc /ō/, /au/ : /o/), except PGmc /i:/, /ī/, which are kept apart in Gothic by a difference in height.

If Marchand, and others who claim that vowel length was non-phonemic in Gothic, are right, then Gothic must

have undergone a quantity shift similar to the Scandinavian one, only approximately a thousand years earlier. The results were, however, much more drastic for the vowel inventory of Gothic than for the Nordic languages in that only in the case of /i:/ /ī/ was the distinction taken over by a qualitative difference. We have, as far as I can see, no means of deciding for Gothic just how this change came about, for example whether long vowels were shortened in front of consonant clusters and short ones lengthened in front of a single consonant and a hiatus, in which case the change was originally context determined, or whether it was context free, simply a loss of a feature, which, incidentally, must have had a considerable functional load in the language.

When these considerations are borne in mind, it seems reasonable to consider whether the synchronic evidence can be taken to be as conclusive as Marchand takes it to be.

Recently, Vennemann (1971) has made a case for length as being distinctive in Gothic. He points out that synchronic processes indicate that a distinction was made between the historically long and short /i:/ /ī/ and /u:/ /ū/. He points out that the breaking before h, h, r, which seems to have been an active allophonic rule in Gothic, only affects the reflexes of short PGmc /i/ and /u/. That the breaking only affected reflexes of historically short /i/, but not reflexes of historically long /ī/, is shown for example by alterna-

tions in the strong verb gateihan (1. class), with the 1. pers. past plural gataihum as compared to greipan, which belongs to the same inflectional class and shows the 1. pers. past plural gripum. The breaking only occurs in the 1. pers. past plural of gateihan, where a historically short /i/ precedes /h/, but ei, representing a historical long /ī/ is left unchanged. That the historically short and long u were also kept apart with respect to the breaking is shown by examples like brukjan 'use' (inf.) (with a historically long vowel) bruhta (past) without breaking as opposed to bugjan 'buy' (inf.) (with a historically short vowel) bauhta (past) with breaking in the past before a h. The last example furthermore suggests that the alternation was synchronically active in Gothic, since u and au alternate in morphologically determined environments. Whenever the h appears after a historically short u, it is turned into what is represented in the spelling by au, probably phonetically something like [ɔ]. This, as Vennemann points out, indicates that the reflexes of short and long PGmc. u, as well as i, were phonologically different in Gothic, but it does not show that the distinctive feature in either case was length.

That the feature that kept the vowels apart was indeed length is, according to Vennemann, shown by the different behavior of these phonemes with respect to Sievers' law (Vennemann 1971:106-109). The so-called Sievers' law is a peculiar behavior of the inflectional

endings in the so-called ja-stem nouns and verbs. When a ja-noun has a long or a polysyllabic stem the nominative ending is -eis, but when the stem is short, the ending is -jis: hairdeis 'shepherd', vs. harjis 'army'. A similar distribution prevails in the ja-verbs: waurkeis 'you work' vs. nasjis 'you save'; the ending after a long root is -eis, but after a short root it is -jis. Vennemann points out that the term 'long' in this connection comprises roots of the form VCC... and also roots of the form $\bar{V}C$, where \bar{V} represents a historically long vowel: sokeis 'you look for' and veneis 'you hope' have -eis: (the comparative evidence points in both cases to a historically long vowel, cf. OI sœkir (with i-umlaut $\bar{o} > \bar{œ}$) 'you fetch, go after' and vân, f. 'hope' (with PGmc $\bar{e} > \bar{a}$)). The roots sok- and ven- can only be long if they have a long vowel, and if the long vowels are considered to be bimoric, the stems sok- and ven- will have the same number of moræ (3) as the VCC stems, i.e. VVC. The high vowels behave in the same way as \bar{e} and \bar{o} in this respect: when ei precedes a single consonant the ending is -eis and when i precedes, it is -jis: gasleibeis 'you damage' vs. bidjis 'you ask', and brukeis 'you use' (with a historically long vowel) vs. hugjis 'you think' (with a historically short vowel) show the same for u.

This evidence for the distinctive feature being length is, however, not as strong as Vennemann maintains. The argument hinges on the assumption that Sievers' law

was a synchronically active phonological process in Gothic, but this cannot be taken for granted. The other Germanic dialects, as well as Gothic, show reflexes of Sievers' law. Old Icelandic can be considered to have developed two different inflectional classes, the so-called ia-stems (old long ja-stems) like hirðir (gen. hirðis) 'shepherd' and ja-stems (old short ja-stems) like herr (gen. hers) 'army'. It is quite impossible to incorporate Sievers' law into OI phonology as an active phonological process. The same applies to Old High German and Old English; they show reflexes of Sievers' law in their inflectional system, but can hardly be taken to contain it in their synchronic phonology. The fact that Sievers' law has left marks in all the other Germanic dialects must be taken, by the comparative method, as evidence that it operated in Common Germanic or Proto Germanic. If this is so, one cannot exclude the possibility that it was fossilised in Gothic as it is in the other dialects. So the fact that the ending -eis appears after historically long roots but -jis after historically short ones in Gothic proves nothing for the synchronic phonology of Gothic. Indeed, Venne-
mann himself mentions examples from the morphology of Gothic, which he calls exceptions to a synchronically active Sievers' law in Gothic. These are neuter ja-stem, and masculine jan-stem nouns. These have the same ending in the genitive, regardless of whether the stem is long or short, whereas, as we saw, the masculine ja-

stems and the ja-verbs show distribution according to Sievers' law. Both kuni 'kind, kin' with a short stem and arbi 'heritage' with a long stem which are ja-stem neuter nouns have genitives in -jis: kunjis, arbis, and similarly wilja 'will' and bandja 'captive' which are jan-stem nouns with a short and long stem respectively, both have genitives in -jins. There seems to be no way of accounting for this exception except by morphological features, as Venemann (op. cit:110) does. But this seems to indicate that the Sievers' law alternations in Gothic were morphologically rather than phonologically motivated, and in that case the length distinction in vowels had nothing to do with the synchronic reflexes in Gothic of Sievers' law, which probably was phonological only in Proto-Germanic or Proto-Indo-European.

The fact that Sievers' law does not prove anything as to whether length was phonemic in Gothic vowels does not mean that it wasn't. I find it just as likely that length distinguished between ei and i and long and short u. The different behaviour of historically long and short u with respect to breaking shows that they were different, whether that difference was in quantity or quality (or perhaps both). It seems, then, that the spelling did not make a distinction prevalent in the phonology of Gothic, namely that between historically long and short u. What this distinction was, is of course hard to say, but it seems at least as likely as not that it was length. It is a well known fact that

length, especially in vowels, was irregularly represented in Germanic writing (cf. e.g. Benediktsson 1968 and Keller 1908) and this could well have been the case in Gothic. If we know for a fact that PGmc /u/ and /ū/ were distinct in Gothic and that the orthography did not reflect that distinction, we can put the question like this: is it more likely that quantity distinctions were left unmarked in the orthography than quality distinctions? This is probably not a very easy question to answer, but given the fact that quantity was not marked consistently in Germanic manuscripts and that it was not generally marked in Latin orthography (cf. Allen 1965:64-65), even though it was distinctive, it seems to be very likely that this was the case in Gothic too, and the difference between the two u-s was that of length. It seems then more likely that the difference in ei:i and the two u-s was length and that e and o were also phonologically long even though their only short counterparts were conditional allophones of i and u respectively. The length opposition is also likely to have distinguished between a as in dags 'day' and a (an) as in fahan 'get', even though nasality cannot be excluded.

To conclude this section on Gothic, let me say this: it seems likely that length was distinctive in Gothic (a) on historical and comparative grounds, (b) because breaking seems to be sensitive to a distinction between two u-s, not marked in the orthography and (c) the distinction most likely to be left out in the spelling is

length. However, this is far from being proven, and can perhaps never be.

The main problem with Gothic is of course that it has left no descendants (except Crimean Gothic) among the modern Germanic languages, so we are missing an important piece of evidence that can be used in the case of the other dialects, namely comparative evidence from younger stages of the language. And this means also that we don't have a history of Gothic to compare with the developments in the other Germanic languages, which might have given important clues.

6. German

It would go beyond the limits of this thesis to account for the development of quantity in German in any detail, but a brief survey, by way of comparison with the Scandinavian phenomena, is in order.

Old High German and Old Saxon had distinctive vowel length, and, for Old High German at least, there were no distributional limits on combinations of long and short vowels with following long (geminated) and short (single) consonants or consonant clusters. Old High German could thus have stressed syllables of the three types we have set up for Old Norse, namely short (short vowel + one consonant), long (long vowel + one consonant or short vowel + long consonant or consonant cluster) and overlong (long vowel + long consonant or consonant cluster). But if we look at the situation

in the modern dialects we see that there is a marked difference between the Scandinavian languages (except Danish) and Modern German, in that in most German dialects, vowel length is undoubtedly phonemic. However, the old, we may call it Germanic, quantity system has not been left intact in German dialects. All German dialects reported on in Keller (1961) show some traces of changes which are reminiscent of the Scandinavian quantity shift. Two historical rules of a quantity shift type are particularly widespread, applying with most regularity in the North. These are a lengthening of vowels in open syllables:

$$V \rightarrow V: / \text{ — CV}$$

and a shortening of long vowels in front of two or more consonants:

$$V: \rightarrow V / \text{ — } C_2$$

(Cf. e.g. Paul/Mitzka 1963:77-79, Von Kienle 1960:37-42.)

The open syllable lengthening shows traces in all German dialects except the southernmost ones (Schwyzer-tütsch (High Allemannic), cf. Von Kienle 1960:37, and Keller 1961:44, 93-94). It seems to have originated in the northern part of the German dialect area and is reported (Von Kienle 1960:37) to have been active in West Low Franconian already in Old High German times, i.e. before 1050. The shortening of vowels before consonant clusters seems to have been more irregular, in that different clusters shorten the preceding

vowels at different times (Von Kienle 1960:40-41), but in most modern dialects vowels are short before two or more consonants.

The North Saxon dialect around ^{the} Lower Elbe (near Hamburg) described by Keller (1961:339-381) shows very clearly reflexes of these changes. Both the open syllable lengthening and the shortening before consonant clusters have taken place regularly in this dialect, but (as in Danish and many other German dialects) old short vowels remain short in monosyllables (closed syllables). This means that vowel length does not become completely predictable by the following consonantism, but there are still regular morphophonemic alternations between long and short vowels within inflectional paradigms, for example where there was formerly an alternation between mono- and bisyllabic forms. Thus, the sg. of Dag 'day' and Slag 'blow' have short vowels, being historically monosyllabic, whereas the plurals, which are originally (underlyingly?) bisyllabic, have long vowels: Daag 'Slääg'. Also the verbs greipen 'to seize' and legen 'to tell a lie' have long vowels in their infinitives, but in the 3rd. pers. sg. pres., where the ending -t is added, forming a cluster following the vowels, the vowels are short (Keller 1961:349). It is striking how similar these phenomena are to the Danish situation described above, and it comes as no surprise, given the geographical proximity and the cultural relations between the areas in question.

As mentioned above, the open syllable lengthening and the shortening of vowels before consonants have not reached all German dialects, especially the southern ones. In these there are, however, traces of other quantity shift like changes. Among these is the so-called Leichtschlusdehnung, according to which vowels are lengthened in monosyllables ending in 'lenis' (cf. Keller 1961:45-48) consonants. In Upper Austrian (Keller 1961:203-218) the combination of this and the lengthening of vowels in open syllables (here called Leichtinnendehnung) has led to a system where quantity is predictable according to the following consonantism: "Every vowel before a lenis or a nasal plus lenis is long, every vowel before a fortis or nasal plus fortis is short" (Keller 1961:204). Even though the results in Austrian are reminiscent of the results in Scandinavian as far as the predictability of vowel length is concerned, this must not be overemphasized, since the historical changes that led to this similar situation in both cases are different, and indeed the environments governing the vowel length are different, being in Scandinavian the length of the following consonantism, but in Austrian the quality of the consonants. (The fortis/lenis distinction, I take it, can not simply be said to be a length distinction.)

Before leaving German, I would like to comment briefly on the development of the consonants. As mentioned above (Section 4.), Danish does not have any

geminated consonants, having undergone a consonant degemination. We have also mentioned that German (the northernmost dialects at least) has undergone a similar change, there being no opposition between long and short consonants. We suggested that this degemination affected the development of quantity in Danish and could (partly at least) account for the unscandinavianness of Danish with respect to quantity. There is very scanty mention of the German degemination in the handbooks I have consulted, but a look at Keller's description of the dialects shows it clearly. In Danish, this degemination has been connected with the weakening of medial stop consonants (klusilsvækkelse) and other phenomena, which had minimized the functional load of the long/short distinction (cf. Rasmussen 1972:67). In many central German dialects a weakening of medial consonants, similar to the Danish one, has taken place (cf. Mitzka 1954). We see, then, that Danish is unscandinavian in more respects than having a distinctive vowel length, and it is an interesting question how these things are related, whether the degemination can have had some influence on the development of vowel quantity, and whether the degemination (which, incidentally, seems to have taken place in English too) is a Danish innovation or whether it spread from the West-Germanic dialects, or perhaps the other way around. But in order to be able to answer these questions, one would have to take a close look at the chronology and geographical distribution of these

phenomena, and there is no room for such an investigation in this context.

7. English

The development of quantity in English has been treated as a whole in Lass (1974), and most of what follows will be a recapitulation of that. Other works dealing with quantity in English from the historical point of view are for example Vachek (1959), Dobson (1962) and Grundt (1973).

In Anglo-Saxon, length was phonemic in vowels (as well as consonants), and the general rule was, as in the other oldest Germanic dialects, that long and short vowels could occur in any stressed environment. It is maintained by Vachek (1959:446) that the length was basically gemination, or 'bimoricness' (cf. Ch. V, Sec.1 below). The only exception to the principle that long and short vowels had a free distribution is that in final open stressed syllables only long vowels occurred. This was caused (Lass 1974:326) by a lengthening of stressed word final vowels, which goes back to Common West-Germanic (if not Common Germanic) times.

As was the case with all the other (surviving) Germanic dialects, this system suffered a series of blows, which led to, or aimed at (cf. Lass), the disruption of the 'Germanic' quantity structure. These changes are listed by Lass (1974:327-333) as the following:

1. Shortening of long vowels in front of sequences of three consonants (6th-7th century).
2. A shortening of long antepenultimate vowels before two consonants. (6th-7th century.)
3. A lengthening of vowels before clusters of liquid or nasal plus homorganic voiced stop. (Around the end of the 9th century.)

In the 11th century, generalizations were made of the two 6th-7th century shortenings mentioned above:

4. A shortening of long vowels before sequences of two (instead of the earlier three) consonants.
5. A shortening of long vowels in antepenultimate position in front of only one consonant (instead of two as before).
6. The last common English thing to happen was the so-called 'open syllable lengthening', according to which vowels (particularly non-high ones) lengthened in the first syllable of bisyllabic words with one consonant following. Along with or before (as a prerequisite for?) the lengthening, the high and mid vowels lowered. These changes took place in the 12th to 13th centuries.

These changes all contributed to making vowel length predictable in an increasing number of environments. In fact, the only places where it was free after these changes was in monosyllables ending in single consonants. But this was enough to maintain a dichotomy in the system

between phonologically different vowels, which derived from the old long (diphthongal) and short vowels respectively. This dichotomy was reinforced by a later neutralization in some of the environments that had come to determine the length of vowels. These were a shortening (degemination) of long consonants (Jespersen 1909/1961:146) and the loss of the final 'weak e' (ibid.:186-189). These changes, which probably took place in the 14th and 15th centuries, removed two sets of environments on the basis of which vowel length was predictable. The degemination removed shortening environments (long consonants) and the loss of the final e removed lengthening environments by turning bisyllabics into monosyllables and thus closing formerly open syllables.

One need hardly emphasize the similarity of these English changes to the development in Danish and German. As a consequence of these changes the quantity situation is similar in all these three languages and different from that of the Scandinavian ones (apart from Danish). There are certain environments where the length (or 'tenseness') of vowels is predictable, but others where it is not, and thus the length of the vowels is still phonemic.

Although in most English dialects vowel length (or 'tenseness') is thus 'phonemic', or free, there is one important exception in that Scots has developed a system where vowel length is predictable to a great

extent. This was brought about by changes that took place in the 17th century, according to which all long vowels and diphthongs shortened everywhere except in front of voiced continuants (r, v, z, ð) and a boundary, and the non-high short vowels e a o lengthened in the same environment (i.e. where long vowels stayed long) (cf. Lass 1974:320). This change, which has come to be called Aitken's law since its exposition by Aitken (1962), led to a situation in most Modern Scots dialects where vowels (except the reflexes of Middle English i and u) have long and short allophones according to the environment: long before r, v, z, ð and a boundary and short elsewhere. The exceptionality of the ME high /i u/ seems to be that they were not affected by the lengthening that occurred in front of r, v, z, ð and a boundary. Thus the vowels [ɛ] and [ʌ] that are the reflexes of ME i and u in the Fife dialect of Modern Scots (cf. Lass 1974:316) only appear as short, whereas other vowels, as a general rule, have both long and short variants.

Lass points out that Aitken's law can be seen as the (nearly) last step in a series of directed changes... that seem to aim at making vowel length predictable on the basis of the environment. (Op. cit.:326) He mentions that the Scottish situation is reminiscent of the Scandinavian one since in most Scandinavian dialects vowel length can be predicted on the basis of the following consonantism. Here, there are also exceptions, as we

have seen, and one might then want to say that Scots is on a par with those Scandinavian dialects that have 'almost made it' to predictable vowel length.

I think, however, that the similarity should not be overemphasized. We have seen that in Upper Austrian, vowel length has become predictable by the following consonantism. But I pointed out that both the environments of the rule governing the length and the historical changes that brought about this situation in Austrian were different from the corresponding Scandinavian phenomena. Similarly, it is important, I think, that Aitken's law, both viewed as a historical change and a synchronic rule of length distribution is quite different from the things we have seen from the Scandinavian languages. Perhaps the most important difference between Aitken's law and the Scandinavian changes is that the central change in Scots is, according to Aitken (1962), a general shortening of long vowels (with the above mentioned exceptions), whereas the Scandinavian quantity shift seems to have 'aimed at' producing 'long' syllables, by lengthening short vowels in front of no more than one consonant (cf. Less 1974: 335). Scandinavian vowels are only shortened if they occur in front of a long consonantism. Thus, whereas Scandinavian (apart from Danish) has now as a rule only long syllables, Scots sides with German and English in having both 'short' and 'long syllables', cf. e.g. [di:v] 'deafen' vs. [dif] 'deaf' (Aitken 1962:2), the

length of which is determined by the length of the vowel. Of course, the importance one assigns to this difference between Scots (and Upper Austrian) and the Scandinavian languages will depend on what one sees as the most important feature of the Scandinavian system. If one emphasizes the predictability of vowel length, one would perhaps not consider the difference as far as syllable types are concerned to be too significant, but if one sees the syllable structure resulting from the Scandinavian quantity shift as its most important feature, one would not want to assign much significance to the similarity between Scandinavian and Scots. The analysis of quantity in Icelandic proposed in Chapter II above emphasizes the syllabic nature of quantity and its relation to stress, and in Chapter IV it will be proposed that rules of stress and syllable shape were the forces that basically determined the development of quantity in Icelandic. If these ideas are justified, (and they can also be applied to Norwegian and Swedish), it follows that there is a basic difference between the development of quantity in Icelandic, Norwegian and Swedish on the one side, and the rest of the Germanic languages on the other. The elimination of free vocalic quality was, as it were, much more tentative and the striving for unity in syllable structure was much weaker in the West-Germanic dialects and Danish than in the northern dialects.

Although it seems to me that there is thus an important split in the Germanic dialect area concerning the development of quantity, it must of course not be forgotten that all the surviving Germanic dialects show a tendency to eliminate free length in vowels and, thus, in a larger context, one can say that the differences between the development in the North and in the South and West are merely variations on a common theme. Indeed, the 'dephonologization' of quantity is not confined to the Germanic languages; a similar thing happened in Latin and Greek. Sommerfelt (1951/1962) emphasizes that the development of quantity and stress in the languages of Western Europe has a common core: "the function of energy takes the form of stress, and the quantitative differences, where they subsist, are entirely subordinate to stress." (p.83.) Sommerfelt suggests that this common development started in Greek and gradually spread to the Western European languages. But if this is the outline of the explanation of the theme, we have still to explain the variations that the different languages and dialects select.

Our question of whether Scots can be said to be Scandinavian in its treatment of quantity turns into the question of whether the Scottish variation was similar to the Nordic one. As I have already said, I think that Scots should be said to have used a different method from the Scandinavian dialects, but it must be admitted that Scots has almost gained the same results as the

Scandinavian languages in making vowel length predictable, and of course the same can be said about Upper Austrian (still in a slightly different way).

Chapter IV.

THE DEVELOPMENT IN ICELANDIC.

0 Introductory

0.1 The prosodic and segmental aspects of the problem

When dealing with the problem of quantity in Icelandic, it is important to bear in mind that it has two aspects, which one should keep apart in theory, even though they are obviously related. On the one hand, one can look at the problem from the point of view of the segmental system and the paradigmatic relations of phonemes and distinctive features. For Old Icelandic, for example, a vowel system can be set up with a distinctive feature of length as a central one, distinguishing between two more or less parallel subsystems (see e.g. Benediktsson 1959:286-295 and 1972:137-138, 146) whereas Modern Icelandic,^{as} shown above, has lost this feature as a distinctive one, and the distinctive function carried by length has been taken over by quality features. We can then say that the quantity shift was a substitution of the length feature by quality features, being brought about by a series of phonological changes like lengthening of vowels in some environments and shortening in others. The taking over of the distinctive function by quality features was

made possible by a group of changes or quality shifts in the vowels. We can call this the paradigmatic or distinctive feature/segmental aspect of the change in so far as we are looking at the effect of the change on the inventory of distinctive features and segmental phonemes and their relations.

But there is another aspect of the problem, which is equally important and which may even prove to be the more important one when we start looking for explanations as to why and how this change came about. We may call this the suprasegmental or prosodic aspect of the problem. From the 'prosodic' point of view the quantity shift is not a loss of length from the phonological system, but we may say that it is reflected in the fact that Old Icelandic and Modern Icelandic have different prosodic structure.¹⁾ Old Icelandic allowed, theoretically at least, for four combinations of vowels and consonants under stress:

1. A short vowel plus one consonant: fat 'a piece of clothing'
2. A short vowel plus two or more consonants: fatt 'erect' (neuter)
3. A long vowel plus one consonant: fát 'confusion'
4. A long vowel plus two or more consonants: fátt 'few' (neuter)

Two mathematically possible alternatives are missing from the table, namely those of a long or a short vowel without a following consonant. When a vowel appears in this environment, in other words in front of another

(syllabic) vowel or a morphological boundary, as in snúa 'to turn', fé 'money' it has been suggested by Benediktsson (1968:40) that the length distinction was (from the paradigmatic point of view) neutralized, since there was no opposition between long and short vowels there. An older theory is that vowels were long in this environment, and in that case, a phonotactic constraint, preventing short vowels from occurring in front of a hiatus or a boundary, must have prevailed in Old Icelandic. Disregarding for the moment the question of the length of stressed vowels in front of other vowels or boundaries, we can make the following statement about the difference in prosodic structure between Old Icelandic and Modern Icelandic. Of the four possibilities of syllable length shown above for Old Icelandic stressed syllables, only two appear in Modern Icelandic, that of a long vowel plus one consonant and that of a short vowel plus two or more consonants. This means that the types 1. and 4. have disappeared, and one can say, as we have seen, that all stressed syllables in Modern Icelandic are in some sense of the same length. One can then describe the difference between Old Icelandic and Modern Icelandic from the prosodic point of view as being that in Old Icelandic stressed syllables varied in length, whereas in Modern Icelandic they are all of the same length. Similarly, we can say that the quantity shift consisted in eliminating syllables of type 1. and 4.

0.2 Sources of evidence

The sources of evidence about the phonology of older stages of Icelandic have been described by Benediktsson (1972:116-117) as being mainly the following:

1) The orthography of written texts can give valuable information about the phonological structure and sometimes even phonetic properties of the language it represents.

2) The metrics of poetry from different times in the history of Icelandic. In the case of quantity, our attention inevitably centers on the rhythmic structure of the metres, because the rhythmic rules of metres can give us valuable clues about the prosodic nature of the language on which the metre is constructed. Of particular interest here is the dróttkvætt-metre, which, it is reasonable to believe, based its rhythm, in part at least, on length variation.

3) Comparative evidence, that is, evidence based on what we know about stages of Icelandic, other than any particular one under investigation, and also evidence based on what we know about related languages. In our case it is, for example, important to know that all the other Germanic dialects seem to have had, at their earliest stages at least, distinctive length in vowels, from which it can be inferred by the comparative method that length as a distinctive feature is a Germanic inheritance and that, at some stage, Old or Prehistoric Icelandic had this feature. Indeed, comparison with

Other Indo-European languages shows that a length distinction in vowels can be traced back to Proto-Indo-European times.

4) For Old Icelandic of the 12th century, there is a very important source of evidence in addition to the three mentioned above. This is the so-called First Grammatical Treatise, which was written in the 12th century and is mainly intended to suggest a solution to the problem of adapting the Latin alphabet to Old Icelandic. This, of course, gives invaluable evidence about the phonology of 12th century Icelandic, and all the more so because it is a remarkable piece of linguistics. (See Benediktsson 1972 for an edition of the text and a thorough commentary.)

1. The phonology of Icelandic about 1200

1.1 The vowel system

We will start our investigation into the history of quantity in Icelandic by summarizing what is known with relative certainty about the phonological system of Icelandic around the end of the 12th century and the beginning of the 13th century.

According to Benediktsson (1959, 1965, 1972) the vowel system shortly after the time of writing of the First Grammatical Treatise (sometime between 1125 and 1175) was dichotomous, divided in two by the feature length. Following Benediktsson we can set up the

Following vowel system for Old Icelandic around or shortly after 1200:

Short			Long			Diphthongs		
i	y	u	i:	y:	u:	ei	ey	au
e	ø	o	e:	ø:	o:			
a	ǫ		æ:		ǫ:			

The symbols can be interpreted roughly as the IPA symbols. The only symbol needing explanation is 'ǫ', which stands for a vowel (historically derived from ǫ by u-umlaut), which probably was a low, back or central vowel, distinguished from /a/ by rounding.

As we can already see from the diagrams, it is assumed that there was not a one-to-one correspondence between the long and the short subsystems. It is only the nonlow vowels which can be said to have a regular correspondence between long and short: /i/-/i:/; /e/-/e:/; /y/-/y:/; /ø/-/ø:/; /u/-/u:/; and /o/-/o:/. In the low vowels, we have probably an opposition of roundness in the short vowels, /a/ vs. /ǫ/, whereas in the long subsystem the opposition between the two low vowels was probably that of frontness: /æ:/ vs. /ǫ:/. The main reason for assuming this difference in the hierarchy and function of the features in the two subsystems is the subsequent development. In the short vowels /ǫ/ shortly after this merged with /ø/, the result being a front vowel, usually represented by the symbol /ø/ (phonetically in Modern Icelandic [œ]). The argument is that if /ǫ/ were a primarily back vowel, it would

have been less likely to merge with a front rounded vowel rather than a back one like /o/. In the long system, on the other hand, a merger took place between /æ:/ (i-umlaut of /ɑ:/) and /ø:/ (i-umlaut of /o:/), two front vowels, differing in roundness, the result probably being a front unrounded vowel. In this case, it can be argued that it shows a relative stability of the backness-frontness feature that /æ:/ merged with /ø:/, retaining its frontness, rather than merging with /ɑ:/, which presumably was kept apart from it by the back-front feature. These arguments may not be conclusive, but in^{the} absence of any arguments invalidating or contradicting the ones presented above we may assume that the relation of the features of backness-frontness and roundness in the low vowels were not the same in the long and the short subsystems (cf. Benediktsson 1959:287-295).

Excursus:

It is worth pointing out that the analysis of the vowel system described above and worked out by Benediktsson is purely surface phonemic. There is for example no attention paid to morphophonemic alternations between vowels, which are, at this stage, quite regular, for example between /u/ and /y/, /o/ and /ø/ and /a/ and /e/ as a result of the historical i- (or i~~R~~)-umlaut, as flytia-flutti 'move' (present vs. past), kómur-koma 'come' (2. pers. sg. vs. infinitive) telja-talða 'count' (present vs. past).

Also, reflexes of the historical u-umlaut show up as regular morphophonemic alternations between /a/ and /ø/ as in kalla-kollum 'call' (inf. vs. 1. pers. plural), bern-born 'child' (nom. sg. vs. nom. pl.) Similarly, in the long vowels, regular alternations resulting from the i-umlaut show up between /u:/ and /y:/, /o:/ and /ø:/, and /a:/ and /æ:/, as in súpa-sýpr 'drink' (inf. vs. 2. p. sg.), fór-fóri 'go' (past indicative vs. past subjunctive) and hár-hæri 'high' (positive vs. comparative). This might lead generative phonologists to suggest that i-umlaut and u-umlaut are active phonological processes in Icelandic of around 1200 and that the underlying phonemic system can be simplified accordingly, assigning for example [y] and [u] to the same underlying systematic phoneme /u/, and [o] and [ø] to the same underlying systematic phoneme /o/, and the [e] of telja to an underlying phoneme /a/ etc. for the i-umlaut, and, similarly, for the u-umlaut, deriving [ø] from underlying /a/ by a u-umlaut rule. This is, of course, a question that deserves careful attention, and I will not set out on a lengthy discussion on the matter, but I would like to make a few points, which seem to me to speak against such an analysis (see S. Anderson 1974:141-146 and Cathey and Demers 1976 for proposals of generative analyses of this sort.)

Firstly, in many cases, the i's and u's which

would cause the umlauts do not appear on the surface, as for example in koma-kømr or brók-brókr for i-umlaut and in barn-born for the u-umlaut. This would mean that in order for the umlaut-rules to be statable in a simple way, the generative phonologist would have to set up abstract systematic phonemes which don't appear on the surface in the relevant positions and would have to be exterminated by special rules. Furthermore, this would not go for all the umlaut-causing segments, since some of them seem to appear on the surface, as for example the u in kollum. The fact that there is not an if-and-only-if relation between for example ø and a following u seems to me to indicate that the process is not phonological, since the environment for ø is not definable in phonological terms except by setting up abstract entities which have no justification on the surface except the umlauted sound itself. This is still more evident in the case of the i-umlaut, since here there are forms which have unumlauted sounds in front of i as in tali 'talk, speech' (dat. sg.). This, I think, shows clearly that the i-umlaut is not a phonological process, definable in terms of phonetically motivated phonological features, but rather a morphophonemic one, conditioned by inflectional categories. It may turn out that the morphology should contain some statements about the morphological function of the

alternations /u/-/y/ etc., but that, I think, should be called morphophonemics, not phonology, and in what follows, it will be assumed that the surface phonemic system described above should be preferred to a more abstract one like:

i	u	i:	u:
e	o	e:	o:
a			ɑ:

It may be added that we would have to make some special arrangements to account for the different results of i-umlaut in the long and the short subsystems. The i-umlaut of long /ɑ:/ would have to give a front low vowel [æ:], whereas the i-umlaut of short /a/ will have to give a front mid vowel [e]. Either there will have to be two i-umlaut rules, one for the long vowels and another for the short ones or a special mechanism of some sort is needed to raise the outcome of the i-umlaut rule, when applied to /a/, from *[æ] to [e]. An attempt could be made to justify an automatic raising of *[æ] to [e] to eliminate a low front vowel in the short subsystem, since there is no low back vowel either, and therefore the feature [±back] has no place in the short low vowels. - We have already suggested that the short /a/ was probably a central vowel. The long surface phoneme /ɑ:/ seems to have been a back vowel, on the other hand. - If some justification of this sort could be found for the raising of *[æ] ,

this latter objection to an abstract analysis is perhaps not so strong, but notice that the motivation for a raising of *[æ] to [e] is sought in the symmetry of the system, and what is more, that symmetry is the symmetry of the surface phonemic system. The raising rule would have the task of eliminating *[æ], which would spoil the surface phonemic symmetry. There is a circularity in this, as I hope is obvious. A justification for a complexity in the rule mechanism of a generative analysis is sought in the regularity of a surface phonemic system (systematic phonetic representation, in terms of generative phonology), which, according to the generative theory as put forward by Chomsky and Halle (1968), has no significant status, either in the linguistic system or in the descriptive mechanism.

(End of excursus.)

In the system described above, the feature of length seems to have played a central role, but it is important to note that it can not have had equal functional status in all the vowels. In the nonlow vowels it seems reasonable to assume that the main difference between the corresponding long and short vowels was indeed length, that is, /e/ and /e:/, for example, had approximately the same phonetic qualities apart from length. The same can be said with reasonable certainty about the pairs /i/-/i:/, /y/-/y:/, /u/-/u:/, and /o/-/o:/.

In the low vowels, however, we have no such one-to-one correspondence. There was no long counterpart to short /e/, and there was no short counterpart to long /æ:/. It is furthermore likely, as we saw above, that in addition to differing in length, /a/ and /ɑ:/ also differed phonetically with respect to frontness-backness. It is more difficult to say anything about the exact phonetic relation between long /ɛ:/ and short /ɛ/, but in view of their subsequent development, it is quite likely that their phonetic properties, apart from length, were different, since the long /ɛ:/ lost its roundness but the short one retained it in most environments. In the cases of /e/, /ɛ/, /ɛ:/, /a/, /ɑ:/, and /æ:/, then, it is quite possible that the function of length as a distinctive factor had already become less important than in the other vowels by about 1200.

In a recent paper (Garnes 1975b) it has been suggested that quantity had already been replaced in the 12th century as a distinctive feature by differences in quality, the long vowels having diphthongized. This gives us a good reason to evaluate the arguments that can be put forward for or against vocalic quantity being distinctive around 1200.

As for the orthographic evidence, there are three old manuscripts which show a regular marking of what has usually been assumed to be length in the vowels. These are Stock. Perg. 4^o No. 15, 'The Book of Homilies' from about 1200, NRA 52., fragments of the oldest saga

of Ólafur Haraldsson, from the first half of the 13th century, and GKS 2087 4^o, the Annales Regii. These manuscripts use an acute accent mark to distinguish historically long vowels from historically short ones. A thorough investigation of the evidence of the Book of Homilies was made by Benediktsson (1968), and his conclusion is that the accent mark was used mainly to mark length, and that irregularity in the notation of forms where the vowel preceded an internal or an external word-boundary was caused by the fact that quantity was neutralized in this position at the time of writing of the manuscript. Benediktsson's investigation also shows that the accent mark was quite frequently used over digraphs denoting the diphthongs /ei/, /ey/, and /au/, its occurrence on these digraphs ranging (according to the different hands) from 24.8% to 75.3% in front of consonants and from 0.0% to 30.0% in front of hiatus or a boundary. These data lead Garnes to propose that the accent mark was used in the Book of Homilies, and also in the later manuscripts, NRA 52 and GKS 2087 4^o, not to denote difference in length, but rather diphthongal quality. There are a number of queries one can make concerning this hypothesis.

It is true that the old long vowels /a:/, /o:/, /e:/ and /æ:/ (< /ø:/, /æ:/) became diphthongs in the history of Icelandic. Evidence for this is to be found for /e:/ as early as in the 13th century in sporadic spellings like ei or ie in place of the older e or é spellings.

Around 1400 the regular notation for the old /e:/ was either e or ie (Þórólfsson 1925: XIV-XV and 1929b:233-234). This was interpreted by Benediktsson (1959:298) as showing that /e:/ probably tended quite early to diphthongize to [ei] , but in order to avoid a merger with the old diphthong /ei/, it subsequently turned towards [ie] and later became [je]. The Modern Icelandic reflex of this vowel is [jɛ] which is best analysed as a sequence of two phonemes, /j/+/ɛ/. Evidence for the diphthongization of /æ:/ to its Modern Icelandic reflex [ai] is not to be found until about 1400, according to Þórólfsson (1925:XVIII). It is important to note that this is only indirect evidence and nothing but a terminus ante quem for the diphthongization. The evidence is spellings like dæginn for older daginn. This form has the Modern Icelandic reflex [ðaijɪn], the main change that took place being a palatalization of the velar fricative represented in the spelling by g. The fact that the symbol g is used in denoting a sequence, phonetically something like [aj], shows that it must have, by 1400, come to represent diphthongal quality in forms like sækja. This could have been going on for quite some time before the sequence [ay] became [aj], which accidentally caused the diphthongization to show signs in the spelling.

Indirect evidence of a different sort can be put forth to indicate that old /a:/ had started to diphthongize quite early. Old short non-high vowels in front of

a nasal + a velar (ng/nk in the spelling) show diphthongal reflexes in most Modern Icelandic dialects, OI langur, MI [laun̥gʏr], 'long', (masc.). The old high short vowels /i/ and /u/ show ^{here} the MI reflexes [i] and [u]: OI ungur, MI [un̥gʏr] (instead of *[Yn̥gʏr]) 'young', OI þing, MI [eiŋ̥] 'parliament' instead of *[eiŋ̥]. This has often been taken to indicate that a lengthening took place in these environments before the quantity shift. It is just as likely, however, that the change before ng/nk was a diphthongization, the vowels developing a high glide in front of the velarized nasal. This is suggested by the fact that the reflex of short /ø/ (< /ø/, /ø/) is usually represented in manuscripts that show signs of the change in front of ng/nk by au in these environments, which is the regular symbol for the old diphthong /au/, MI [œy]. Signs of this change before ng/nk show up in the manuscripts as early as around 1300 in that the old short vowel /a/, for example, is represented by symbols, which otherwise denote the old long /ɑ:/. If the change in the old short /a/ before ng/nk was a diphthongization, then the usage of the symbols previously only used for the old long /ɑ:/ would seem to indicate that it had begun to diphthongize as early as about 1300. (This was pointed out to me in correspondence by Stefán Karlsson of the Arnamagnæan Institute in Reykjavík.) The high vowels pose no problem for the hypothesis that the change before ng/nk was a diphthongization, since the glide that was added was

a high one agreeing in roundness with the original vowel, and if indeed the change was not a simple lengthening, but rather a raising of the last part of the vowel, it is only natural that the MI outcome should be fully high vowels /i/, /u/, rather than /I/ and /Y/.

It may be added that it is usually assumed that the diphthongization of the old long vowels must have preceded the quantity shift, both because the diphthongal quality had to be there, when length was neutralized, to take over its function in the system, and because the environment for diphthongization is usually assumed to be length, and the phonetic diphthongization, it is assumed, must have been completed before the old long vowels started developing shortened allophones.

It is, then, consistent with Garnes' proposal that there must have been a period before the quantity shift was completed, when the old long vowels /e:/, /æ:/, /o:/, and /ɑ:/ were diphthongized, and in that respect it is conceivable that the accent marks over these vowels denoted diphthongal quality in the early manuscripts. But there is another fact that does not fit her proposal, namely that the high vowels /i:/, /y:/ and /u:/ are still monophthongs in Modern Icelandic, and it is highly unlikely that they were ever anything else. Thus, if Garnes were right in assuming that the accent mark denoted diphthongal quality, it should have been left out over the high vowels. This is, however, not the case. It is impossible to interpret the statistics

adduced by Benediktsson (1968) as showing any significant difference between the occurrence of the accent mark over the high and the non-high vowels.

Another fact which speaks against Garnes' proposal is that the First Grammarian explicitly states that there is a length correlation in the vowels. The passage is as follows in Benediktsson's translation:

.... But even though I do not write more vowel symbols than the vowels that have been found in our language - eighteen made out of the five Latin vowels - it is well to know that there is yet another distinction (greiⁿ, KÁ) in the vowels - both in those that were in the alphabet before, and in those that have now been put in - a distinction which changes the discourse, (according to) whether a letter is long or short, just as the Greeks write a long letter with one shape, and a short one with another. Short g they write this way: ε, but the long one like this letter is: η; short o in this way: ο, but the long one in this way: ω. This distinction, too, I wish to show, because it changes the discourse just like the previous ones, and (I shall) mark the long ones with a stroke (to distinguish them) from the short: far: fár, rám:rámr, ól: ól

(Benediktsson 1972:218-221)

The most natural interpretation of this is to take it as if the First Grammarian is talking about a quantitative correlation between otherwise similar vowels, and we would have to have very good grounds for interpreting this in any other way. Garnes tries to cast doubt on the validity of this evidence by arguing that the First Grammarian was 'hard-pressed to come up with minimal pairs' to show opposition between long and short vowels (Garnes 1975b:4). She mentions that a number of the examples the First Grammarian gives are not minimal in the strictest 20th century sense. We must note, however, that most of the suspicious examples are not ones dealing solely with vowel length. Three examples pertain to a length distinction in consonants, namely y be, the names of the letters y and b as opposed to Ubbe, a man's name, ho do 'a tall (woman) died' (two words) as opposed to hoddo, gen. sg. of hadda 'handle' (one word) and afarar as opposed to afferar. (It is difficult to say exactly what the words in the last pair are, since the sentence which presumably followed, illustrating the meaning, is left out in the only extant manuscript of the treatise.) (Benediktsson 1972:244-245). These three pairs are, as I said before, intended by the First Grammarian to illustrate the difference between long and short consonants and have nothing to do with the question whether length was distinctive in the vowels. Two other examples which Garnes cites are intended to illustrate opposition between nasality and non-nasality, namely i sá 'one

could see through' (literally 'in saw') with a nasalized long i as opposed to isa, acc. pl. of is 'ice', and bu at 'you (were) at' as opposed to þuat 'pressed down' with a nasalized ú (Benediktsson 1972:218-219). The examples which have to do with vowel length are the following: sebo vs. sébo, framer vs. fra mér, and Gobróbi vs. gób róbi. The first member of the first pair is probably imperative of the verb to see, séa plus the 2nd pers. sg. pronoun þú 'thou'. The second member is assumed to^{be} past 3. pers. plur. of the (irregular) verb sýia 'nail together'. In this case the vowel of sé, which is historically long, is opposed to another historically long one. This is the only example which can be said to cause problems, since identical vowels seem to be used to illustrate a difference in length. The next pair, framer vs. fra mér is intended to illustrate the opposition between a short and a long nasalized e. True, the -mer in framer 'forward, brazen' (pl. masc.) is a second syllable of a bisyllabic word, and therefore presumably unstressed, whereas mér is an independent word, 'me' (dative), but there is no doubt that the two e-vowels were distinct. What probably forced the First Grammarian to use such a far-fetched minimal pair was the difficulty of finding a long and^a/short nasalized vowel in minimal opposition. The third pair, gób róbi vs. Gobróbi is even less problematic. This pair is intended to illustrate the distinction between a long /ø:/ and a short /ø/. The first member of the pair is two words, gób, neuter

plur. of the adjective góbr 'good' plus róbi, 'oars', whereas the second is a compound man's name. Here again, there is no real problem; -róbi and róbi were undoubtedly kept apart by short vs. long vowels. There may also have been some differences in stress, one being a second part of a compound but the other an independent word, but there is no doubt that the vowels were distinct. It will be noticed that there is also a distinction in length in the first parts of the examples, that is, between gób, and Gob- of Gobróbi. In this sense, the pair is not minimal, but we notice that it is non-minimal in a special way, since it is in both vowels length that distinguishes, and it may well have seemed appropriate to the First Grammarian to ~~throw~~ in this extra example of the length distinction as a decoration. We can perhaps call Gobróbi vs. gób róbi a double minimal pair.

To return briefly to the only problematic example, namely sé bú vs. sébu, Benediktsson has (1972:138-139, cf. 1968:42-44) proposed an explanation, according to which the opposition long vs. short was neutralized before a hiatus and a boundary. Benediktsson also proposes that the 'archiphoneme' occurring in the neutralizing environment was identified with the short vowels at the time of writing of the First Grammatical Treatise, and therefore the First Grammarian could use sé with an e-sound in neutralized position to illustrate a short vowel opposing a long one. If this explanation is valid, there is no problem. There seems, however, to be some-

thing not quite right about assuming that a nondistinctive occurrence of a feature (that is, the feature occurring in environments in which it is redundant) can be used to illustrate the distinctive function of that feature. There may be a way out of this problem. We may very well say that the quantity distinction was neutralized in front of a hiatus and a boundary, since no minimal pairs with a short vowel opposing a long one in this environment are found, but it is interesting to see what happens when there is a morphological alternation between forms ending in a vowel and forms with a consonant following. There are many examples of this to be found, for example in nouns with a stem ending in a vowel taking consonantal endings: tré 'tree' genitive trés. A similar alternation is to be found in weak verbs like fá 'paint' when they take the preterite ending -ba or the past participle ending -b, as in fába, fáb 'painted'. Phonologically we seem to have a forced choice between a distinctively long or a short vowel, unless the morphological boundary played a major role in phonology, which seems rather unlikely, especially in the monosyllabic forms trés and fáb. In cases like these we seem, then, to have a morphophonemic alternation between a neither-long-nor-short 'archiphoneme' and a long or a short vowel. In these cases the choice between a long or a short correlate of the archiphoneme could go either way, that is, we could either have an alternation: 'archiphoneme' -short vowel or an alternation: 'archi-

phoneme'-long vowel. In sebu, written in one word in the manuscript, we could easily be dealing with such a case. As we saw before, the form consists of the stem of the verb to see plus an enclitic pronoun -bo (or bu) . In Modern Icelandic the corresponding form is sjáðu, where the -ðu can be analyzed as an ending (cf. Árnason 1974: 28) even though it is historically an enclitic pronoun. The same can have been true of the form sebo; the -bo may have behaved like a regular inflectional ending from the phonological point of view, forcing a choice between a long or a short vowel, and in this particular paradigm the alternation could have become 'archiphoneme' - short vowel. If this was so, then the e-sound in sebo was a perfectly legitimate short vowel and could be used to demonstrate the distinctive function of length in the e-vowels.

Finally, it can be pointed out in connection with the First Grammarian's evidence and Garnes' proposal, that she does not maintain that the distinction between historically long and short vowels had disappeared, rather that length had been replaced by features of quality as distinguishing marks between the old long-short vowel pairs. In that case, the First Grammarian's 'difficulty' in finding minimal pairs would be just as difficult for Garnes to explain as it would be for those who maintain that length was the distinctive feature. Even if the function previously held by the length feature had been taken over by some qualitative differences

at the time of writing of the First Grammatical Treatise, the vowels were still kept apart and The First Grammarian should have had no trouble in finding pairs that showed the distinction. The paradox is that Garnes agrees that the vowels were distinct, but claims that the First Grammarian had difficulty in finding minimal pairs because the distinctive feature(s) was (were) qualitative instead of quantitative before.

To summarize then, the testimony both of the Book of Homilies and of the First Grammatical Treatise seem to agree on a dichotomy which corresponds to that between historically long and short vowels. Furthermore, the First Grammarian explicitly calls this distinction one of length. From our review of the evidence, we have so far found no reason to disbelieve him.

There is still one fact which must not be overlooked, even though it is often overlooked, and which Garnes seems to overlook, perhaps because it is so obvious. This is the fact that the Icelandic orthography, which was formed in the 12th century, uses the same symbols to denote long and short corresponding vowels. How can that be explained? The most likely explanation is that they must have been phonologically related. As in Latin writing the main purpose of the symbols was probably to denote vowel quality, and it seems to be hard to explain why the same symbols were used for /i/ and /i:/, /e/ and /e:/, /o/ and /o:/ and /a/ and /a:/ except to assume that they had approximately the same quality but were distinguished by length.

1.2 The prosodic system

1.2.1 The evidence of metrics.

As was pointed out at the beginning of this chapter there are two aspects of the length problem in Icelandic, (and of course the other Scandinavian, and indeed Germanic, languages), the paradigmatic one concerning the vowel system and its distinctive features and phonemes, and the prosodic one, concerning the rhythmic structure of the language. We will now turn our attention to this latter aspect.

As has been said many times, there were four types of combinations of vowels plus consonants theoretically possible at the stages of Icelandic which still had distinctive vowel length. These were: VC, V:C, VC₂, V:C₂ (C₂ denotes two or more consonants, including geminates, analysed as two identical consonants.)

From the prosodic point of view it is customary to classify these four types of syllables according to length, and then a syllable consisting of a long vowel + a consonant is grouped together with a syllable consisting of a short vowel + a long (geminated) consonant or two or more consonants, these being called long. The syllables consisting of a long vowel + a long consonant or two or more consonants are called overlong. The syllables of the type VC are called short. From the prosodic point of view, the diphthongs are assumed to have functioned as long vowels. The long vowels and diphthongs can be considered to have had two moræ,

whereas a short vowel had only one. We thus get the following classification of syllables: 2)

Short:	VC	<u>fat</u> 'a piece of clothing'
Long:	V:C	<u>fát</u> 'confusion'
	VC ₂	<u>fatt</u> 'erect' (neuter)
Overlongs:	V:C ₂	<u>fátt</u> 'few' (neuter)

As a source of evidence for features of the sort just mentioned a prime candidate must of course be metrical rhythm, and, naturally, in the literature on quantity in Icelandic, metrics of poetry from different times is commonly called upon as evidence. (See e.g. Þórólfsson 1929a, Karlsson 1964 and Benediktsson 1968.) It is, however, as pointed out by Benediktsson (1968: 46-47), important to bear in mind that metrics cannot be taken without comment as direct evidence about linguistic facts, since the metres have rules of their own, and there is not necessarily a one-to-one correspondence between linguistic features, such as for example stress or length, and the rules of metrics, even though they are evidently related. The linguistic system is not the same as the metrical system, although the latter is based on the former.

Early modern writers on Old Icelandic metrics, such as Sievers (1893), noted a correlation between quantity and stress on the linguistic side and the scansion of poetic text into more or less regular feet (pedes, Füsse) on the metric side. But Sievers and others noted discrepancies between e.g. the linguistic quantity on the

one side and its value in the metre on the other. It seems, for instance, that monosyllabic forms of the type VC (a short vowel plus one consonant) could function both as 'Hebungen' (ictuses) and 'Senkungen' (drops) (Sievers 1893:58) in the Eddic metres.

If we are right in assuming that Old Icelandic had stressed syllables of varying length, that means that this length variation could be used to create a rhythm based on regular alternations between long and short syllables, as was done in classical Latin poetry. We must remember, however, that in all likelihood length was not the only rhythmic or prosodic feature in the language which could be utilized to create poetic rhythm. There is no doubt that stress was also a very important feature, and it has indeed been considered to be one of the most influential ones in the history of the Germanic languages, both as far as prosody is concerned (see e.g. Sommerfelt 1951/1962:82-83) and in the effects it had, directly and indirectly, on the paradigmatic relations in the phonological systems. We can then say that Old Icelandic provided two means for poets to create rhythm in their verse : length and stress. They could both alternate long and short syllables and stressed and unstressed ones. When we study the rhythmic laws of Old Icelandic poetry, we must, then, consider both possibilities. It is possible that one metre based its rhythm on stress alternations, and another on length alternations, and we cannot exclude the possibility that some or all metres used a mixture of both.

Old Icelandic poetry is usually divided into two different types, the so-called Eddic Poetry and the so-called Skaldic Poetry. The former is usually considered to be of older origin than the latter, the metres and much of the subject matter being of common Germanic origin. The skaldic poetry, on the other hand, is considered to be purely Nordic in origin, although it is often considered to be partly due to Celtic influence. It is the poetry of the (mostly Icelandic) skalds (skáld, 'poets') who were often employed at the courts of Norwegian kings and made poems about the heroic ventures of the kings. These poems were often recited at the courts for the entertainment of the kings and their warriors. The metres of these poems were much more rigid than the Eddic metres, both rhythmically and as far as rhyme and various other poetic devices are concerned.

The most important mediaeval authority on skaldic poetry is Snorri Sturluson in his Háttatal 'Inventory of Metres', which forms one part of his Edda, which was most likely intended as a handbook for poets who wanted to keep up the old tradition of skaldic poetry. Háttatal is considered to have been completed during the winter 1222-23. It is a poem about the Norwegian king Hákon Hákonarson and his protector, earl Skúli. The poem is so composed as to show the various metres that could be used in skaldic poetry and had been used, according to Snorri, in the skaldic tradition since its

start in the 9th century. To go with the poem, Snorri wrote a commentary on the metrics, explaining the peculiarities of each metre, and these commentaries are interspersed between the stanzas illustrating the metres. As a rule, one stanza illustrates each metrical variant.

The most common of the skaldic metres and no doubt the most original and basic one was the so-called drótt-
kvaður háttur or dróttkvætt (meaning originally 'the court metre') which consisted of 8 lines (vísuorð) to each verse. Each line consisted, according to Snorri, of 6 syllables. The verses were further decorated and bound together by internal (to the line) rhyme and alliteration. As an illustration of the metre Snorri gives the following stanza:

Látr, sá's Hákon heitir
hann rekkir lið, bannat
jörð kann frelsa, fyrðum
friðrofs, konungr, ofsa;
sjalfr ræðr allt ok Elfar
ungr stillir sá, milli,
gramr á gipt at fremri
gandvíkr, jöfurr, landi.

(Cf. Jónsson 1912-15 AII:52 and BII:61)

What we are particularly interested in is the rhythm. As we see, Snorri is consistent in having six syllables to a line in this verse, but later he says:³⁾

"It is permitted that the metres have syllables slow (seinar) or quick (skiótar) so that there is

an increase or a decrease from the correct number of [i.e. of syllables according to] the rules, and [syllables] may be found so slow that five syllables are in the second and the fourth line [of each half verse], as is here:

7. Hjalms fylli spekr hilmir
hvatr Vindhlés skatna
hann kná hjörvi þunnum
hras þjóðár rassa,
ýgr hilmir lætr eiga
öld dreyrfá skjöldu,
stýrs rýör stillir hersum
sterkr járngrá serki"

(Snorra Edda 1931:218. My translation and italics.)

It is evident that Snorri's basic metrical unit is the syllable (samstofun), but he makes an interesting distinction between types of syllables, slow and quick, and it is very unlikely that he is speaking of anything but our distinction between long and short syllables. If we look at the lines 2, 4, 6 and 8 of the 7th stanza of Háttatal which Snorri refers to as being exceptional as they only have five syllables, we see that all syllables except the final (unstressed) ones are either of the type VC₂, V:0 or V:#. If we, for example, scan the second line with the symbols —, denoting a long syllable and ∪, denoting a short one, we get:

hvatr Vindhlés skatna

We see that all the syllables, except the very last one, are long according to our classification above. Similarly, we can scan the fourth line in the following way:

hræs þjóðár ræsa

and the 6th and 8th lines:

öld dreyrfá skjöldu

sterkr járnrá serki

The last two lines call for a minor footnote. They both have second parts of compounds -fá and -grá, ending in a vowel, which we have here scanned as long. But, as will be remembered, Benediktsson considers vowels in final position to be phonemically neither long nor short, and about 1200 largely 'identified' with short vowels. If the vowels were phonetically short, the syllables -fá and -grá should perhaps not be metrically long. But in the first place, we are not sure that they were phonetically short in every occurrence, and in the second place, in these two forms we have special cases of word final á. The forms are both plural accusatives of adjectives having stems ending in long /a:/, their nominatives singular being, respectively fár and grár (the -r is a nominative ending). The accusative plural masculine ending for adjectives is -a, and the underlying forms of fá and grá, as they appear in the context above are probably /fa: + a/ and /gra: a/, and the final vowels of the surface forms could in these cases well have been phonetically long and identified with long vowels, even

though otherwise word final vowels were identified with the short phonemes. In any case, it is quite common for syllables of this type to carry the ictus in skaldic poetry (cf. Benediktsson 1968:42-43).

If we look at the other lines of the 7th stanza viz. 1, 3, 5 and 7, which Snorri seems to consider more regular, we see that they all have two short syllables, that is, the final one of the line and one other, appearing somewhere in the middle of the line. In the first line, we have fyllí, in the third hjórví, in the fifth hilmir, and in the seventh stillir, all supplying a short syllable, making the total number of syllables Snorri's regular six. This seems to indicate that Snorri had some notion of a rule about how many long and short syllables could occur in one line. If the number was below six, all but the last syllable had to be long.

A look at the next stanza of Háttatal confirms the hypothesis that length of syllables played a role in the dróttkvætt-metre. To this verse Snorri gives the following introduction (again in my translation): 4)

"Now, there shall be shown syllables, so quick and put so close to each other, that the length of the line is increased because of it:

8. Klofinn spyr ek hjalm fyrir hilmis
hjarar egg, duga seggir;
því eru heldr, þar er skekr skjöldu,
skafin sverð lituð ferðar;
bila muna gramr, þó at gumna

gular rítr nái líta,
draga þorir hann yfir hreinna
hvatan brand þromu randa "

(Snorra Edda 1931:218)

I would like to start the analysis of this stanza by looking at lines 2, 4, 6 and 8, because they have a remarkably regular pattern. They all have seven syllables, and if we scan them according to the length of the syllables, we get:

2 hjarar egg duga seggir
4 skafin sverð lituð ferðar
6 gular rítr nái líta
8 hvatan brand þromu randa

If we define each short syllable as having metrically one mora and each long syllable as having two, we notice that in these lines, we get the same number of moræ as in the corresponding lines of verse 7, that is 9 moræ. This clearly shows that short and long syllables, as we defined them above, had different functions in the dróttkvætt-metre; two short syllables could equal one long one metrically. (This is evidently the same thing as happens in Greek, Latin, Old English and other Germanic verse, called 'resolution'.) This, of course, does not amount to saying that in regular speech every long syllable was twice as long as a short one. We must remember to keep the linguistic system and the metrical system apart. But this gives considerable support to our hypothesis that in Old Icelandic around

1200, there was, prosodically, a difference between long and short syllables. This evidence can then be added to the evidence concerning the paradigmatic distinction between long and short vowels.

The odd numbered lines of verse number 8 seem to be more irregular, and it can be said that they are not very typical dróttkvætt-lines, far less so than the even numbered ones. As Snorri says, they contain nine syllables, if everything is counted, but it is very likely that in lines 1, 3 and 5 at least there are examples of liaison of two syllables into one (bragarmál in Snorri's terminology). Spvr ek of line 1, for example, is probably to be scanned as one syllable, spvr'k, and bar ex (or ea) of line 3 is to be scanned bar'a, and hó at of line 5 is to be scanned as hó't. Thus, in these lines, the number of syllables (and more) is cut down by one. But there are still more more than the nine of the even numbered lines. Snorri's comment about syllables being put close to each other must apply specially to these lines, and what he has in mind is either the above mentioned liaison or the sort of skidding over the syllables one often hears in singing, when the words don't quite fit the rhythm of the music they are sung by. Anyway, what interests us here is the fact that apart from the syllables carrying the internal rhyme (rhyming with the last but one syllable, which is always long according to the rules of the dróttkvætt-metre) the syllables are short. There are two syllables

other than the ones mentioned above which could perhaps be taken as long, namely skekr and bó't (< bó at). However, these forms probably functioned metrically as short syllables for reasons we need not go into here.

The conclusion to be drawn from this brief look at Snorri's ideas about the rhythm of the dróttkvætt-metre and his practice in using it, is that there is a clear distinction between long and short syllables and that they had different functions in the metre. This must have been based on some distinction in the prosodic structure of the language, and we may take this as evidence that Old Icelandic had, at the time when Snorri composed his Háttatal, at least two types of syllables, long ones and short ones. Another fact which has long been noted in writings on the dróttkvætt-metre confirms this. This is that it seems to be an exceptionless rule that the last foot (the last two syllables) of every line must consist of a long syllable followed by a short (unstressed) one. Thus, a line of a dróttkvætt-verse could only end in forms like skatna (VCC), bunnum (VCC), ressa (V:C), eiga (V:C or VVC), réttum (VCC) etc. As far as I know, there are no examples of dróttkvætt-lines ending in forms like daga, tala, sonar (all VC) etc. This rule could hardly be upheld unless there was linguistically a clear distinction between short and long syllables.

When this metrical testimony is added to the evidence given by the orthography and the First Grammatical

Treatise there can hardly be any doubt that a length difference prevailed in the vowels, since a historically long vowel plus one consonant was treated prosodically as long, whereas a short vowel plus one consonant was treated prosodically as short.

1.2.2 Overlong syllables

Before leaving the prosodic aspect of length in Old Icelandic, we will have a brief look at the hypothetical distinction made above between long and overlong syllables. We have seen that the evidence of the metrics of skaldic poetry, more specifically that of dróttkvætt as described by Snorri Sturluson in the first quarter of the 13th century, confirms the distinction between long and short syllables. When it comes to the overlong syllables, however, there is no evidence in the poetry that they had a function different from the regular long syllables. Overlong syllables carried the last ictus in the line, that is, we get rétum, traustar, hárr etc. (all with V:CC or VVCC) alongside sennu, leita, gáti etc. as the last foot of a line, and I have not been able to find any other signs of their special metrical status in a survey of 1438 dróttkvætt lines from the 10th to 14th centuries. We can't automatically say that this shows that there was no prosodic difference in the language between long and overlong stressed syllables. Even though there was such a difference in the language, it would not necessarily mean that it had to

show up in the rules of the metres used in poetry. But on the other hand it is quite conceivable that these syllables, having underlying long vowels and original diphthongs followed by two or more consonants were in fact phonetically no longer or not significantly longer than the other long syllables. In other words, it is quite possible that the underlying long vowels had shorter allophones when they were followed by two or more consonants than when they were followed by one or none. Whether or not an allophonic rule, shortening long vowels in front of two or more consonants, already existed around 1200, it is certain that at some stage such a rule must have arisen. Something of this sort must have been the historical ancestor of the shortening part of the length rule in Modern Icelandic.

It is important to note, however, that if such a rule was operative at an early stage, it did not lead to a large scale restructuring in lexical items, for example so that the shortened long vowels merged with the phonemically short ones. There are only scattered examples of this in words like gott (with a Modern Icelandic monophthong /ɔ/: [gɔht^h] which is the regular reflex of Old Icelandic short /o/), the neuter of góður with an originally long root vowel. Similarly, Modern Icelandic drottning 'queen', and drottinn 'king, master', both with an [ɔ]-vowel, presuppose a pre-vowel-shift shortening of the /o:/ in dróttning and dróttinn (both derived from drótt 'court, army'). Further, the Modern

Icelandic nominative minn of the first person possessive pronoun 'mine' (with |I|, the MI reflex of OI short /i/) as opposed to the dative minum (with MI [i], the regular reflex of OI /i:/) shows that OI /i:/ must have been shortened before the two consonants in the nominative. The main rule, however, is that the old long vowels show the same reflexes in Modern Icelandic as far as quality is concerned, regardless of whether they precede one or two or more consonants: hvítur [kfi:t^hYr] and hvítt [kfiht^h] 'white' (masc. vs. neuter) both have vowels with [i]-quality as descendants of old long /i:/.

Even though it does not seem that a shortening of long vowels in front of two or more consonants, making them merge with old short ones, was a regular phonological change that took place before the quantity shift proper, it is still conceivable that such a low level phonological rule existed quite early. It could have operated without leading to a merger of shortened long vowels with the original short ones, just as soon as there appeared qualitative differences between the corresponding members of the long and short subsystems. As soon as, for example, the short vowels began to lower, as they must have done at some stage, giving e.g. the modern [ɔ] as a reflex of the old short /o/, and the non-high long phonemes began to diphthongize, giving e.g. modern [ou] as the reflex of the long /o:/, it was possible to shorten old long /o:/ before two or more

consonants without a general merger with old short /o/. It is perhaps not even necessary to assume a great qualitative difference between the two subsystems for this to be possible. As long as a speaker could unambiguously identify the (underlying) phonological identity of each allophone, a relatively minor phonetic difference would be needed to keep allophones of different phonemes apart (cf. Grundt 1973:139 and passim). Clues of different sorts, other than surface phonetic ones, may help the speaker to establish the underlying phonological origin of a surface sound. Clues of this sort may be the syntagmatic phonological surroundings of the phone in question or, perhaps even more important, the morphological identity of the form that the sound appears in. It is conceivable that a speaker identifies the [o]-sound of for example the form skjótt, 'quickly' relatively similar to the [o]-sound of skot 'a shot' with underlying /o:/ because of the evident morphological relationship with skjótur 'quick' (masc.) and skjót id., fem. which had both underlying and superficial long o-sounds. In terms of features, one can picture this by setting up a special phonetic feature [shortened]. The surface [o]-sound of skjótt could then be characterized as [+long, +shortened] with the feature [+shortened] added by a phonological rule, whereas the similar [o]-sound of skot can be characterized as merely [-long], and [+long, +shortened] and [-long] may turn out to represent more or less the same

phonetic reality, but the underlying origin of the two phones may be easily recoverable by phonological, morphological or even semantic means. A phonological clue could have been that the [o]-sound of skiótt appears in shortening environments for a long vowel, a morphological clue could have been that skiótt is related to skiót and skiótur, and a semantic clue could have been that skiótt has some of the same semantic features as skiótur and skiót.

It seems, then, quite conceivable that the old long vowels were allophonically shortened in certain environments before the quantity shift proper had taken place, that is, before length was replaced by qualitative features as distinctive between the old short and long vowels. There is, however, one thing which could make it difficult in our case to assume that the long vowels had short and long allophones according to environment before the qualitative features took over the function previously held by the length feature. As we have seen, the Modern Icelandic reflexes of the old long non-high vowels are diphthongs, and these diphthongs appear both in lengthening and shortening environments. OI liótr and liótt give MI [ljou:t^hYr] and [ljouht^h], OI kátr and kátt give MI [kau:t^hYr] and [kauht^h], OI kata and katti give MI [cai:t^ha] and [caihtI], and OI él and éls give MI [jæ:l] and [jæls]. It has long been attested that there seems to be a connection between length and diphthongization. Long vowels show a greater tendency

to diphthongize than short ones. This is true of e.g. Faroese (cf. Rischel 1968 and Árnason 1976) and many modern Norwegian and Swedish dialects and indeed of Modern Icelandic, which shows a tendency to diphthongize the long allophones of the mid vowels /ɛ/, /œ/, and /ɔ/ (Garnes 1974b). Similar tendencies show up in German and English. If we try, then, to find a reason or an explanation for the diphthongization of the Old Icelandic non-high long vowels /æ:/ (< /æ:/, /ø:/), /ɑ:/, /e:/ and /o:/ the most obvious feature to connect it with is the length of the vowels in question. We can say that the length of the vowels created a favorable environment for diphthongization. (We cannot say that length was, or is in general, a sufficient condition for diphthongization, e.g. for the obvious reason that the Old Icelandic high vowels /i:/, /y:/ and /u:/ did not diphthongize.) But it is not enough for us simply to note that it is quite common for long vowels to diphthongize; we would like to be able to explain why this is so. Presumably we would explain it in phonetic terms along the following lines: the longer the duration of a vowel, the greater the chances that an internal variation in the quality might occur, for example by the features of the sound, instead of being in a 'simultaneous syntagm', becoming to some extent temporally ordered (cf. H. Andersen 1972). When we connect diphthongization with length, then, we must mean phonetic length, since in that way we can give some plausibility to explaining diphthongization as a consequence of length.

If we carry this over to the question of allophonic shortening of underlying long vowels, we see that, as soon as the reflexes of the old diphthongs were shortened to any significant degree, the length as a favorable environment for diphthongization had disappeared for the (conceivably) short allophones, and thus, if a rule shortening long phonemes in certain environments became significantly operative before the old long vowels diphthongized, then there seems to be no reason why the α -sound of skiótt should diphthongize. But the fact is that it did. When we look at the phonetic nature of the diphthongization, we seem, then, to be forced to conclude that phonetic diphthongization of the old long vowels had occurred before the shortening of long vowels in front of two or more consonants took place.

It seems, then, even though it was theoretically possible that long vowels in 'overlong syllables' were shortened by an allophonic rule without their merging with old short ones, that the fact that the allophones of the long vowels /ɑ:/, /æ:/, /e:/ and /o:/ before two or more consonants were diphthongized in the same way as the other allophones, indicates that the old long vowels could not have been significantly shortened before two or more consonants until after the phonetic diphthongization had taken place. This would seem to mean that if a shortening of long vowels had taken place around 1200, then the diphthongization had taken place earlier, and in the synchronic grammar, shortening

was ordered after diphthongization, if both were active phonological processes. All of this seems to indicate that if we want to date allophonic shortening as far back as around 1200, we must push the diphthongization even further back in time. But it may be that this is not necessary. We saw above that the phonetic environment for diphthongization must be assumed to have been long duration, and in the discussion that followed, we tacitly assumed that that meant that all and only the phonetically long vowels should diphthongize. But this only follows if we assume that all phonological change has to be explained in phonetic terms and that phonological change is purely additive. It may well be that this is too narrow a view of linguistic change. Under this interpretation, the change is 'Markovian' in the sense that it is assumed that only phonetic surface forms can be referred to in accounting for changes and no sort of underlying systemic relations are taken into account. But it is not necessary to assume that no underlying part of the phonological structure is recoverable. It is conceivable that we have changes that behave like transformations in that they affect and operate on parts of the phonology other than the mere surface forms. Within generative phonology the question of whether phonological change is purely additive, in the sense that only surface forms can change, has often been formulated as the question whether changes only occur in the last rules of the phonology. The general

theoretical framework of generative phonology allows easily for changes in the underlying system and 'early' rules, leaving 'lower' parts of the phonology intact. It is theoretically possible that rules are 'added in the middle of grammars' as the question has been put by King (1974). It is, then, theoretically possible that diphthongization took place in other places than the ones where the phonetic surface conditions, which probably triggered it off in the first place, were present. This could have taken place in the following way: Assume that the Icelandic vowel system had two subsystems, kept apart by a feature which we can call length. There were quality differences between the long and short vowels in addition to the difference in length. The long vowels had shortened allophones in front of two or more consonants. Some of the phonetically long vowels (the non-high ones) started to diphthongize. We assume that a phonetic condition for the diphthongization was the length of the vowels. Phonetically speaking, this environment did not exist in shortened long vowels, but these vowels were qualitatively different from the phonemically short ones, and when the diphthongization was phonemicized, the phonetically shortened long vowels /ɑ:/, /e:/, /o:/ and /æ:/ became underlying diphthongs, because they belonged to the same phonemes as the corresponding long (unshortened) variants.

The account given above of the possible way in which phonological diphthongization historically followed shortening of long vowels in front of two or more consonants hinges on the phrase 'belong to the same phoneme'. It must be assumed that if diphthongization did not take place until after the old long vowels had developed shortened allophones then, when the underlying forms of the still long allophones became diphthongs, the underlying forms of the shortened allophones became diphthongs too. This could only take place if the shortened long vowels still belonged to the same phonemes as the long vowels. 'Belonging to the same phoneme' must mean having the same underlying features, and that brings us back to the feature that divided the old short and long subsystems. We have hitherto called this feature length, but since we have alluded to the possibility that the 'long' vowels had both long and short allophones, one may well ask whether a different term is not suitable. If the feature does not relate directly with phonetic duration in so far as there are allophones of long vowels that are relatively short and others that are relatively long in duration, then why do we call it length? Are we forced to assume that as soon as long and short allophones of the old long vowels started to appear, the feature length must have been replaced in the old long vowels by some other feature or features distinguishing e.g. old /a/ from old /a:/ ? This may turn out to be merely a terminological question as to what we like to call the feature that distinguishes

between the old long and short vowels after duration had become predictable in some environments. It has been suggested by Sigmundsson (1970:321) that the label tense should be used to distinguish between^{the} old long and short systems at this stage. This may at first glance seem a sensible thing to do, but the trouble is that we are no closer to knowing the truth although we invent a new term for the feature. We have already seen that it must be assumed that there was, prosodically, a difference between long and short syllables in the beginning of the 13th century. Syllables like fát were longer than syllables like fat. The length of the former syllable must stem from the length of the vowel, so, whatever name we use for the feature, among its phonetic correlates must be long duration, at least when not followed by two or more consonants. We have also seen that most probably there were differences other than length between the old long vowels and the corresponding short ones, and we now must decide how to distinguish in our model between the old long and short vowels. Whether we call it length or tenseness probably does not make much difference, since in any case there will have to be secondary features derivable from the underlying abstract one. If we call the vowels tense, there will have to be set up a number of rules to predict the length and quality, sometimes diphthongal, sometimes monophthongal etc. If we call the vowels long, the diphthongal quality will

also have to be predicted by some sort of redundancy rules or automatic low level phonetic rules.

Rather than choose one or the other term for this mysterious feature right away (see Chapter V, Section 4 for further discussion of this), I would like to summarize what we can already say about the state of the vowel system and the prosodic structure of Old Icelandic about 1200. The evidence reviewed above strongly speaks against the quantity shift being completed. The First Grammarian explicitly speaks of long and short vowels, and the rules for the rhythm of the dróttkvætt-metre seem to have been defined, to some extent at least, on the basis of a distinction between long and short syllables. This must mean that e.g. syllables like dag(a) vit(a) still remained short at this stage and that the lengthening of the vowel had not yet taken place. There is, however, a possibility that when followed by two or more consonants, the old long vowels had shortened allophones. Yet these shortened allophones, if they existed, did not generally merge with the corresponding short vowels, but stayed allophones of the long vowels. For this to have been possible we must assume that there was not a one-to-one correspondence between the two systems as far as quality is concerned and that there were qualitative differences, though not necessarily great, between the corresponding members of the long and short subsystems. It is likely that already about 1200 or shortly after that there was

a slight difference in quality between e.g. /i/ and /i:/, the short vowel probably being lower than the long one. Similarly, a difference in quality must have prevailed between /o/ and /o:/. The short /o/ was probably somewhat lower than the long /o:/, and it is further quite probable that the long vowel was somewhat diphthongized. We can visualize the relationship between the two sub-systems, after the merger of /β/ and /ϕ/ into /ö/ (around 1200) and of /æ:/ and /ø:/ into /æ:/ (around 1250); in the following way:

<u>FRONT</u>				<u>BACK</u>	
<u>Unround</u>		<u>Round</u>			
<u>Long</u>	<u>Short</u>	<u>Long</u>	<u>Short</u>	<u>Long</u>	<u>Short</u>
/i:/ [i:]	/i/ [ɪ]	/y:/ [y:]	/y/ [Y]	/u:/ [u:]	/u/ [ɯ] High
/e:/ [e ⁱ :] or [e ⁱ es]	/e/ [ɛ]		/ö/ [ø]	/o:/ [o ^u :]	/o/ [ɔ]
/æ:/ [æ ⁱ :]					
<u>Low central unround:</u> /a/ [a]				/ɑ:/ [ɑ ^o :]	Low

In this system, an allophonic alternation between long and shortened long vowels could have existed without a large-scale merger of the long and short vowels in the shortening environments. An allophonic shortening then, did, not have to lead to a restructuring in the lexical phonology. But if we assume that such an allophonic rule existed, we have to think of the (phonological) diphthongization as a rather complicated process. If the diphthongization of the low vowels was not well established and incorporated into the phonological system, but just a

minor phonetic feature, conditioned by the length of the vowels, we would expect the shortened long vowels not to diphthongize, since there, the conditioning factor, viz. long duration, was not present. But if we can assume that the phonological diphthongization consisted in replacing underlying long (but sometimes phonetically shortened) non-high vowels by underlying diphthongs, it is theoretically possible that an allophonic shortening was an active phonological process before the diphthongization became phonologized. In this way we can allow for a shortening of old long vowels (and diphthongs) in front of two or more consonants without at the same time having to assume that the old long vowels had become phonological diphthongs.

I have been trying to show that it is possible that some feature of length still had a major function in the vowel system and was manifested in phonetic duration among other things, even though there were shortened allophones of the long vowels. But of course this does not amount to saying that it was the case that the long vowels had shortened allophones in front of two or more consonants. The only evidence we have about this is that we have found no sign of the 'overlong' syllables having a special metrical function different from the function of the regular long (bimoric) syllables. But this, as has been said many times before, does not prove anything. I find it, however, rather likely that there existed quite early - perhaps even earlier than 1200 -

an allophonic shortening rule for long vowels in front of two or more consonants. There seems to have been a tendency to shorten excessively long syllables. This is shown by the fact that long geminated consonants in front of another consonant are often not distinguished in the spelling from short ones. In early thirteenth century manuscripts, inverse spellings like blindom (majuscules were often used to denote long consonants or geminates) for blindom 'blind' (dative plural) are quite common. This must mean that the difference between long and short /n/ was neutralized in front of a consonant, that is, the forms renndr past participle of renna 'run' and rendr 'rimmed' (derived from the noun rönd 'rim') were homophonous. The most plausible phonetic explanation for that is that the long /n/ was shortened in front of another consonant. In this way, instead of a sequence VCCC we have a sequence VCC, which is the same output from the phonotactic point of view as if we assume that long vowels were shortened in front of two or more consonants.

1.3 Summary. The situation in the early 13th century.

We can now try to summarize what we can say about the situation about 1200 as far as vowel quality and quantity and the prosodic system is concerned. Length seems to have been distinctive in the vowels. 'Prosodically', this meant that there were long and short

syllables in the language. There was also a possibility of 'overlong' syllables. We have not found any evidence that these overlong syllables had any special prosodic or metrical status. It is, then, conceivable that long vowels in front of two or more consonants were allophonically short, making the hypothetical overlong syllables of the same prosodic length as the VCC or V:C syllables. In addition to the difference in quantity between corresponding members of the long and short vowels, there were probably differences in quality. The non-low short vowels were lowered, and the non-high long vowels were slightly diphthongal. It is, then, likely that the feature called length by the First Grammarian was not purely durational, but was related on the phonetic side to qualitative differences as well.

2. From Old to Modern Icelandic

In the preceding section I tried to form an idea of the situation in Old Icelandic about 1200, as far as vowel quantity and quality and the prosodic system is concerned. The conclusion was more or less that there was, prosodically, a distinction between long and short stressed syllables, and that this prosodic difference was based on a paradigmatic distinction between long and short vowels, in that when a syllable consisted of a short vowel plus only one consonant, it was prosodically short. A long vowel plus one consonant, on the other

hand, formed a prosodically long syllable. I also concluded that the difference between the long and short vowels was phonetically not only one of duration, but there were also some qualitative differences between the corresponding long and short vowels. As we have seen, a part of the quantity shift was that the old short stressed vowels lengthened when followed by one consonant. (It is unlikely that vowels were ever phonemically short in forms (like bá, óua 'live' where no consonant followed). We thus have in Modern Icelandic fat [fa:t^h] and fata [fa:t^ha] with long [a]-vowels as reflexes of the old short /a/. In this section we will concern ourselves with the question of how or when the short vowels lengthened and how this lengthening was related to the shortening of long vowels that I discussed in the previous section. The questions with which we will deal can be summarized as the following:

- (1) Did the lengthening occur at the same time in monosyllables and in polysyllabics, that is, did the root vowel of fat lengthen at the same time as that of fata?

- (2) When did the change (or changes) take place?

A third question which we will have occasion to consider in this section is:

- (3) Why and how did the change(s) take place?

This last question is obviously related to the two first ones, and we will have to bear it in mind when dealing

with the questions of (relative) chronology, but we will leave our main discussion of it for Section 3 of this chapter.

2.1 The term 'quantity shift'

In works on the history of Icelandic phonology the quantity shift is usually referred to without comment as if it were one historical change which took place at some (known or unknown) definite point in time. (Cf. e.g. Þórólfsson 1929a, Karlsson 1964, Benediktsson 1959, 1963 and 1968.) The term has been used to refer to the 'disappearance of the quantity correlation' in vowels (Benediktsson 1959:300), and the reduction of the number of syllable types from four to two (of the four types of stressed syllables: VC, V:C, VCC and V:CC... , only two remain: V:C and VCC...) or from three to one (the overlong and the short syllables were eliminated, and now all stressed syllables can be said to be long). I have nowhere seen a comment suggesting that the quantity shift was a complex of changes, which should be dealt with separately. This of course does not mean that everyone believes that it was a single change; in fact there have been very few comments made about the nature of the change in detail.

Regardless of what has been done before, we will obviously have to concern ourselves with this question, that is: Was the quantity shift really a shift that

took place in one step? Our superficial survey of the development in the other Scandinavian languages and our brief look at German and English have shown that changes like the components of the quantity shift don't necessarily have to occur in a block. There are Norwegian dialects which have undergone some components of the quantity shift but not others - Setesdal has retained long vowels and diphthongs in front of two or more consonants, Gudbrandsdal has had no lengthening of short vowels, and the Tinn-dialect has had lengthening only in monosyllables. This shows that each component of the quantity shift can occur without the others occurring at the same time; one could say that each part is a perfectly 'legitimate' sound change. Old long vowels could shorten in front of two consonants without a lengthening of short vowels taking place, and short vowels could lengthen in monosyllables without a lengthening in bisyllabics taking place at the same time. The development in Danish shows that it was also possible that short vowels lengthened in bisyllabics by an 'open syllable lengthening', (for the term, cf. Grundt 1973 and Lass 1974) without a lengthening in monosyllables taking place. (See Weinstock 1975 for a survey of the development of quantity in the Scandinavian languages.) We will obviously have to bear all this in mind when we study the Icelandic quantity shift.

In this connection it is also fruitful to abstract a question of more general nature from the sort of comparative or direct empirical evidence we may have at our disposal and ask, in general, how such a change as the quantity shift could have taken place and wonder, from the point of view of some sort of a theory of linguistic change, what kind of a change it is likely to have been. We can for example ask, apart from considerations of similarity with the development in the other Scandinavian languages and the direct evidence we can produce, whether it is likely that the loss of the quantity correlation in vowels and the prosodic change from three lengths of stressed syllables to one length, took place overnight, so to speak; that is, whether everything happened more or less at the same time. Or is it more likely that a number of changes 'conspired' to give the results mentioned above? (cf. Lass 1974.) It is not a simple thing to answer such a question, because the answer will depend on what sort of a model of linguistic change we believe in, and we are far from being able to say that we have at our disposal a well motivated and explicit theory of linguistic change. If we had a good theory of linguistic change, one way of answering the general question posed above would be to feed the two alternatives, the one assuming that everything happened at the same time and the one assuming that we are dealing with a historical conspiracy, into an evaluation metric which would be incorporated

into our theory of linguistic change. Given the relevant data and the two alternative hypotheses, the evaluation metric should tell us which alternative is 'more highly valued' and therefore more likely to be the correct one. Even though we are not lucky enough to have at our disposal an explicit theory of linguistic change which can automatically tell us which alternative is the better one, there is nothing to prevent us from making our own guesses as to which of the alternative models for the quantity shift, the 'overnight alternative' or the 'conspiracy alternative' is more likely to get higher marks on an evaluation scale incorporated into an adequate theory of linguistic change, (or whether they would perhaps get the same marks, i.e. be equally likely or unlikely, or simply whether the question is undecidable).

In considering the problem of forming an adequate theory of linguistic change, and more specifically of forming an evaluation metric for the naturalness of linguistic changes, at least three kinds of considerations must be brought forth. Firstly, we must consider the extralinguistic context in which language operates; in connection with sound change, we must, for example, consider which model for a particular change is most easily relatable to phonetic reality. In other words, we should form our general theory so as to be able to evaluate the phonetic plausibility of individual models of changes. Another set of considerations which may affect our evaluation measure are considerations of

formalism. We may want to choose one model of a particular change over others on the grounds that the one we choose is more simple to state within the general framework we are working in. We may for example choose one model on the grounds that it requires fewer rules or fewer symbols than others. Criteria of this sort are of course useless unless we can be certain that our formal framework is somehow valid and that simplicity in the description reflects linguistic naturalness or simplicity. This brings us to the third type of criteria which we must consider when discussing evaluation metrics for linguistic changes. This is the bearing that what we know about the structure of language will have on the evaluation metric and consequently on the validity of different hypotheses about particular changes. Obviously, our ideas about language structure will have to be kept as clear as possible of bias from particular descriptive models and formalisms, since in synchronic linguistics the models are only attempts to describe language structure; they are merely put forth hypothetically. But if we can be reasonably certain that we have established some facts about the synchronic structure of language in general (linguistic universals) or about a particular language at some synchronically defined stage, we will want statements which we make about language change to be compatible with these facts, and we can also hope that these facts can help us to make discoveries about linguistic change. In this way we can say that knowledge

about linguistic structure will have ^{an} effect on the way we evaluate hypotheses about linguistic change. In using this sort of evidence in evaluating hypotheses about language changes, we must, of course, make sure that the things we claim to know about the structure of language in general, universal grammar, if you like, or the grammars of particular languages, are relevant to the historical changes we are dealing with. Even though we may know quite a few things about linguistic structure, these things are not necessarily ones that will have a direct bearing on the historical problems we are dealing with. In our case, we may not be allowed to comment on the likely course of events in the quantity shift on the grounds of what we know about the structure of Icelandic before and after the change or on grounds of what we know about linguistic universals, because these facts we know may not be relevant to the problem.

If we now concentrate on the problem facing us, namely how the quantity shift is most likely to have taken place, we have set up a choice between two models, the 'overnight alternative' and the 'conspiracy alternative'⁵⁾. We can start by trying to imagine how we can relate the two alternatives to phonetic reality. We can, for example, ask what phonetic conditions could have triggered the change, or changes, depending on which alternative we (eventually) will choose. We have already suggested that it is likely that the shortening part of the length rule in Modern Icelandic originally

stems from an allophonic shortening of long vowels in front of two or more consonants. In phonetic terms, we can try to make such a hypothesis plausible in the following way: We can assume that a phonetic segment tends to adapt to its surroundings as much as ^{is} allowed by its distinctive function and underlying features. And if we assume that a stressed vowel and the consonantism following it formed some sort of an articulatory unit, perhaps defined by stress (we may like to call this unit a phonetic syllable), it seems, in some way at least, to make sense to expect the duration of the vowel to alternate according to how many consonantal segments followed it, that is, if an underlyingly long vowel was followed by two or more consonants, one could expect it to be relatively shorter than when followed by only one consonant, and shorter than predicted by its underlying features. If we look at the beginning of the shortening of the long vowels in this way, we don't have to assume that it had any connection with a lengthening of vowels in other environments.

Looking at the lengthening of short stressed vowels in front of one consonant and trying to relate it to phonetic reality, we can suggest that when the vowels bore the stress, they tended to become longer than would be directly predicted by their underlying features. This assumption derives some plausibility from the fact that experiments have shown that there seems to be a close connection between stress and duration of vowels

in a great number of languages (cf. Lehiste 1970:125-142 and references). This lengthening of vowels could well take place without the shortening of long vowels in front of two or more consonants taking place at the same time.

There is, then, from the phonetic point of view, no necessity to assume that these two phonetic changes occurred at the same time, and it seems as likely as not that they came up at different times.

There is an important difference between the strength of the claims made by our two alternative hypotheses. The overnight hypothesis makes a stronger claim, in that it assumes that the two changes occurred together. The conspiracy alternative makes a weaker claim. The only thing that is said about the connection between the two changes is that the result they produce together seems to connect them in some way. It is neutral with respect to time, and makes no claims as to whether the changes took place at different times or at the same time. Our look at the possible phonetic conditions for the lengthening of short vowels and the shortening of long ones does not seem to give any support for the stronger claim made by the overnight hypothesis. We can thus say that the conspiracy hypothesis is a better theory of the relation in time between the lengthening of short vowels and the shortening of long ones, since it does not make any unjustified assumptions. In the absence of any evidence to the contrary, it seems, then,

that it is best to assume that there was no connection in time between the lengthening of short vowels and the shortening of long ones. The dating of these two changes is therefore purely a matter of investigation using any external evidence we may find. We are not, of course, saying categorically that the two changes did not occur at the same time, but there is no more reason for us to expect that than to expect them to have taken place, say, six decades apart.

To look briefly at the question of evaluation from the point of view of formalism, let us start by noting that we have as yet no way of telling what formalism is the best or the right one for describing things of this sort, and any formal argument will only be as good as the formalism it is based on. (Cf. Chen 1976, pp. 211 ff. for clear examples where considerations of formal simplicity lead to either wrong or unsatisfactory conclusions in historical linguistics.) We seem to have means of formalizing, at least partly, both the overnight and the conspiracy hypothesis about the relation between lengthening and shortening of stressed vowels. The overnight alternative could be formalized as something similar to a two-sided transformation (with a lengthening and a shortening part) like the one set up in Chapter II pp 8-2 for the synchronic length rule in Modern Icelandic, and the conspiracy alternative could be formalized simply as a set of rules with their chronological order either left unmarked (if we don't

know when the changes took place) or marked by the dates we can put on them using external evidence. I won't set out on what seems to me the hopeless mission of trying to figure out which alternative is formally simpler, if indeed it is possible to evaluate formalisms. And without being able to say which alternative is formally simpler, which would presumably be done with the help of some formal simplicity measure, we are nowhere. In any case, as I have said before, a simplicity or naturalness measure for a formalism will only have value as far as we can show that it fits a certain amount of well established linguistic data. And the sort of data we should be looking for are facts about events like the quantity shift. So, we will find justification for one or the other alternative formalism in facts we can establish about the quantity shift, but not the other way around.

It seems, then, that the most sensible thing for us to do until we find some evidence to the contrary, is to choose the conspiracy alternative, that is, make no assumptions as to the relative chronology of the shortening of old long vowels and the lengthening of old short ones. It may well be true, as I have suggested, that the old short vowels had already shortened about 1200, but something else may also be the truth.

Having considered the relation between the lengthening of short vowels and the shortening of long ones, we may now turn our attention to the lengthening of short

vowels by itself. We have seen that in related languages a lengthening could take place in monosyllables without a lengthening occurring at the same time in polysyllables, and vice versa. It seems, then, a priori equally possible that this was the way things happened in Icelandic. It is theoretically possible that the lengthening took place in two steps. Like the question of the chronological order of the shortening of long vowels with respect to the lengthening of short ones as a whole, this is of course an empirical question. It may be interesting, however, to consider whether there is something in the nature of the lengthening change (or changes) which would make us expect one order rather than the other, that is, whether it is likely that the short vowels lengthened simultaneously in monosyllables and polysyllabics or whether one lengthening preceded the other in time. We have already imagined what the phonetic conditions for the lengthening of the vowels could have been, and we mentioned that it was not unlikely that stress had something to do with it. We have no evidence as to whether the stress pattern of monosyllables and polysyllabics was different in any way in the periods when the lengthening must have taken place. We are perhaps allowed to say that, if anything, we would expect that the monosyllables had more stress concentrated on their only syllable than, for example, the bisyllabics had on the first of their two syllables. We saw (Chapter III, Section 2) that

this sort of thing has been proposed as an explanation of vowel balance and other phenomena in Norwegian dialects. We have absolutely no indication that this was the case in Icelandic, but we can perhaps say that it would not surprise us, if there were any difference in time between the two lengthenings, that the lengthening started earlier in monosyllables. If this were the case, then, the lengthening in Icelandic would be more similar to what happened in Norwegian and Swedish than to what took place in Danish, and in German and English. This would be in conformity with the general fact that Modern Icelandic, Faroese, Norwegian and Swedish all have developed the same prosodic system as far as quantity is concerned.

Although we have more or less excluded formal considerations of rule writing from having any significance, we may mention that it seems, in some sense at least, to be simpler from that point of view to assume that all stressed syllables lengthened at the same time. This is so because it takes up less space on paper and fewer symbols to write down a single rule like:

$$\begin{array}{c} V \longrightarrow V: / \text{ --- } C^1 \\ [+ \text{ stress}] \end{array}$$

which describes a lengthening of all nonlong stressed vowels in front of a single consonant, then to write the two rules:

$$\begin{array}{c} V \longrightarrow V: / \text{ --- } C\# \\ [+ \text{ stress}] \end{array}$$

and

$$\begin{array}{c} V \longrightarrow V: / \text{ --- } CV \\ [+ \text{ stress}] \end{array}$$

with different datings. But it is just as likely as not that this is one more instance where the formalism makes an incorrect prediction.

It remains, then, purely a matter of empirical investigation to find out whether lengthening of vowels was a single historical change (whatever that means) or whether it took place in two steps, separated by some time interval. Both alternatives are conceivable, and it is difficult to say which one is 'more natural' and should be 'more highly valued' in a general theory of linguistic change, and therefore more to be expected.

2.2 Dating the changes

In the foregoing, I set forth some ideas about the sort of internal arguments that could be used, in dealing with the quantity shift, its nature and chronology, that is, I tried to imagine what sort of assumptions we would make about it in the absence of any external, philological evidence. In particular I tried to relate possible hypotheses about the change to phonetic reality. In this section we will have a look at the external evidence we can find. Primarily, this evidence will be given by poetry from different times, more specifically its metrics. This is just about the only evidence we can use, since the orthography of texts, which often can give information about phonological development and can help to date changes, was not affected in any drastic way by the quantity shift,

since there was no change in the phonemic inventory in the sense that phonemes merged or split, and the same graphemic system could be used before and after the change; only the phonetic value of the symbols was changed.⁶⁾

We have seen that, judging from the situation in the Norwegian and Swedish dialects which still have not completed the quantity shift, the short vowels seem to have lengthened earlier in monosyllables than in polysyllables in continental Scandinavian. We saw for example that in some northern Swedish dialects the word corresponding to Standard Swedish väv 'cloth' has a long vowel in the monosyllabic nominative singular but a short vowel in the bisyllabic plural (Chapter III, Section 3.) We also saw that a similar situation prevailed in the Tinn dialect of Norwegian. On the other hand, we saw that in Danish the short vowels only lengthened in polysyllables, old short monosyllables still being short in Danish. There is, then, an interesting difference between the way Danish and the continental Scandinavian languages lengthened short vowels. Danish behaves more like the West-Germanic languages, English and German, and we have hinted at the possibility that German influence could account for the 'unscandinavianess' of Danish in this respect.

In the light of what we know about the other Scandinavian languages, it will now be interesting to see whether we can discover how the Icelandic short-

vowel lengthening proceeded, that is whether it proceeded in the continental Scandinavian way or in the Danish and West-Germanic way, or whether Icelandic took the third alternative, namely lengthened all short stressed vowels at more or less the same time.

It will be clear from what we have already said about the shortening of long vowels that it is most natural to assume that it did not take place at the same time as the lengthening of short ones. Even though we don't have any proof, we have suggested that the shortening took place earlier than the lengthening, perhaps as early as 1200. We will leave the question of shortening and concentrate on the nature and dating of the lengthening of short vowels.

2.2.1 The evidence of metrics

2.2.1.1 Dróttkvætt

We have already seen that the Skaldic metres, in particular dróttkvætt, seem to have had a rhythm which, in part at least, could be defined in terms of rules for alternations of long and short syllables. The metrics of dróttkvætt could then help us to confirm that around 1200, there was a prosodic difference between long and short syllables, the short syllables consisting of an old short vowel followed by no more than one consonant. Of equal importance to us now is the fact that as soon as there was a change in the prosodic structure

of the language, this was liable to show some signs in the poetry. For instance, if old short syllables lengthened, they should become able to take on the function previously only held by old long syllables. We saw for instance that in the dróttkvætt-metre the last but one syllable of each line always had to be long. This is a virtually exceptionless rule in the oldest poetry. The last foot of the line is always a long-stemmed bisyllabic word, i.e. forms like manni, hestum, ásir, nóttu etc.; short-stemmed words like tala, bera, kona, etc. could not stand in this position. Consequently, if we find, at some later time, exceptions to this rule, so that for example tala can occur as the last foot of a line, this must either mean that there has been a change in the rules of the metre or that the first syllable of tala has become long so that it can carry the last ictus of the line.

Quantity played a further role in the dróttkvætt-metre, as we have seen, in that there was at least a tendency to keep the number of morae (cf. above, p. 177) in each line constant, granted that two short syllables could have the same value as one long one (two morae). A very common type of dróttkvætt-line is one that has a regular alternation between long stressed and short and/or unstressed syllables, like this:

— — — — —
undrask qglis landa (Pormóður Kolbrúnarskáld,
lausavisa. Cf. Jónsson
1912-15, AI:288.)

Using the moraic measure, we can scan this line as having 9 mora. It can also be analysed as consisting of three feet, each being a trochee. It is perhaps worth noting that if we assume that the rhythm was based on length alternations, it can be represented in musical notation, assuming that the rhythm is a 3/4 rhythm:

| $\overset{d}{\underset{\cdot}{\text{u}}}$ $\overset{d}{\underset{\cdot}{\text{n}}}$ | $\overset{d}{\underset{\cdot}{\text{q}}}$ $\overset{d}{\underset{\cdot}{\text{l}}}$ | $\overset{d}{\underset{\cdot}{\text{l}}}$ $\overset{d}{\underset{\cdot}{\text{a}}}$ |

The number of this type of dróttkvætt lines is endless.

I will cite here only a few:

$\overset{\sim}{\text{o}}\overset{\sim}{\text{l}}\overset{\sim}{\text{l}}\overset{\sim}{\text{i}}$ $\overset{\sim}{\text{Ó}}\overset{\sim}{\text{l}}\overset{\sim}{\text{a}}\overset{\sim}{\text{f}}\overset{\sim}{\text{r}}$ $\overset{\sim}{\text{f}}\overset{\sim}{\text{a}}\overset{\sim}{\text{l}}\overset{\sim}{\text{l}}\overset{\sim}{\text{i}}$ (Erfidrápa Ólafs helga 1.3)

$\overset{\sim}{\text{m}}\overset{\sim}{\text{i}}\overset{\sim}{\text{l}}\overset{\sim}{\text{d}}\overset{\sim}{\text{i}}\overset{\sim}{\text{n}}\overset{\sim}{\text{g}}\overset{\sim}{\text{s}}$ $\overset{\sim}{\text{m}}\overset{\sim}{\text{á}}\overset{\sim}{\text{l}}\overset{\sim}{\text{s}}$, $\overset{\sim}{\text{e}}\overset{\sim}{\text{n}}$ $\overset{\sim}{\text{g}}\overset{\sim}{\text{u}}\overset{\sim}{\text{l}}\overset{\sim}{\text{d}}\overset{\sim}{\text{u}}$ (ibid. 2.7)

$\overset{\sim}{\text{R}}\overset{\sim}{\text{a}}\overset{\sim}{\text{u}}\overset{\sim}{\text{ð}}$ $\overset{\sim}{\text{i}}$ $\overset{\sim}{\text{r}}\overset{\sim}{\text{e}}\overset{\sim}{\text{k}}\overset{\sim}{\text{k}}\overset{\sim}{\text{a}}$ $\overset{\sim}{\text{b}}\overset{\sim}{\text{l}}\overset{\sim}{\text{o}}\overset{\sim}{\text{o}}\overset{\sim}{\text{i}}$ (ibid. 14.1)

$\overset{\sim}{\text{m}}\overset{\sim}{\text{e}}\overset{\sim}{\text{i}}\overset{\sim}{\text{n}}\overset{\sim}{\text{a}}\overset{\sim}{\text{l}}\overset{\sim}{\text{a}}\overset{\sim}{\text{u}}\overset{\sim}{\text{s}}\overset{\sim}{\text{t}}$ $\overset{\sim}{\text{i}}$ $\overset{\sim}{\text{m}}\overset{\sim}{\text{i}}\overset{\sim}{\text{n}}\overset{\sim}{\text{u}}$ (ibid. 25.3)

$\overset{\sim}{\text{H}}\overset{\sim}{\text{e}}\overset{\sim}{\text{r}}\overset{\sim}{\text{o}}\overset{\sim}{\text{ð}}\overset{\sim}{\text{a}}$ $\overset{\sim}{\text{v}}\overset{\sim}{\text{a}}\overset{\sim}{\text{l}}\overset{\sim}{\text{d}}\overset{\sim}{\text{r}}$ $\overset{\sim}{\text{o}}\overset{\sim}{\text{f}}$ $\overset{\sim}{\text{f}}\overset{\sim}{\text{a}}\overset{\sim}{\text{l}}\overset{\sim}{\text{d}}\overset{\sim}{\text{i}}\overset{\sim}{\text{n}}\overset{\sim}{\text{n}}$ (Hákonarflokkur 2.2)

$\overset{\sim}{\text{R}}\overset{\sim}{\text{i}}\overset{\sim}{\text{k}}\overset{\sim}{\text{r}}$ $\overset{\sim}{\text{g}}\overset{\sim}{\text{a}}\overset{\sim}{\text{f}}$ $\overset{\sim}{\text{h}}\overset{\sim}{\text{l}}\overset{\sim}{\text{e}}\overset{\sim}{\text{n}}\overset{\sim}{\text{n}}\overset{\sim}{\text{a}}$ $\overset{\sim}{\text{h}}\overset{\sim}{\text{n}}\overset{\sim}{\text{e}}\overset{\sim}{\text{y}}\overset{\sim}{\text{k}}\overset{\sim}{\text{i}}\overset{\sim}{\text{r}}$ (ibid. 5.1)

$\overset{\sim}{\text{H}}\overset{\sim}{\text{e}}\overset{\sim}{\text{i}}\overset{\sim}{\text{m}}$ $\overset{\sim}{\text{k}}\overset{\sim}{\text{o}}\overset{\sim}{\text{m}}$ $\overset{\sim}{\text{h}}\overset{\sim}{\text{i}}\overset{\sim}{\text{l}}\overset{\sim}{\text{m}}\overset{\sim}{\text{i}}\overset{\sim}{\text{r}}$ $\overset{\sim}{\text{R}}\overset{\sim}{\text{a}}\overset{\sim}{\text{u}}\overset{\sim}{\text{m}}\overset{\sim}{\text{a}}$ (ibid. 9.5)

(For reference to the texts see below pp.215-16)

Another very common type of dróttkvætt line is the following:

$\overset{\sim}{\text{a}}\overset{\sim}{\text{l}}\overset{\sim}{\text{v}}\overset{\sim}{\text{a}}\overset{\sim}{\text{l}}\overset{\sim}{\text{d}}\overset{\sim}{\text{r}}$ $\overset{\sim}{\text{s}}\overset{\sim}{\text{k}}\overset{\sim}{\text{i}}\overset{\sim}{\text{p}}\overset{\sim}{\text{u}}\overset{\sim}{\text{m}}$ $\overset{\sim}{\text{h}}\overset{\sim}{\text{a}}\overset{\sim}{\text{l}}\overset{\sim}{\text{d}}\overset{\sim}{\text{i}}\overset{\sim}{\text{t}}$ (Hákonarflokkur 1.2)

In this line, instead of having a regular alternation between long and short syllables we have two long syllables in a row followed by two short ones, one after the other. This can also be represented in the following way:

| $\overset{d}{\underset{\cdot}{\text{a}}}$ $\overset{d}{\underset{\cdot}{\text{l}}}$ | $\overset{d}{\underset{\cdot}{\text{s}}}$ $\overset{d}{\underset{\cdot}{\text{k}}}$ | $\overset{d}{\underset{\cdot}{\text{h}}}$ $\overset{d}{\underset{\cdot}{\text{a}}}$ | $\overset{d}{\underset{\cdot}{\text{l}}}$ $\overset{d}{\underset{\cdot}{\text{d}}}$ |

If we count the mora of this line, we see that we still

have 9, only now they are distributed differently over the words in the line. Further examples of this type of scansion are:

tállaustr við þála . (Erfidrápa Ólafs helga 1.2)
sókn stríðs fyrum ríða (ibid. 1.6)
úthlaupum gram kaupast (ibid. 4.2)
hann sýfði svá þýfðir (ibid. 5.4)
rauð brúnan hjör túnun (ibid. 14.8)
auðmíldr sáker gildi (Hákonarflokkur 2.4)
Uplond farit bröndum (ibid. 3.6)
víkelds gisfir ríkum (ibid. 9.2)

A third type of line, which is also very common, we saw in the 8th stanza of Snorri's Háttatal:

hjárar egg daga seggir etc. (Cf. p.177)

Here, instead of the line starting with a long syllable, it starts with two short ones, but the number of more is still the same. Further examples are:

snerir fundast þar þrænda (Erfidrápa Ólafs helga 11.3)
jöfur magnar guð fegas (ibid. 25.2)
sumir skunduðu undan (Hákonarflokkur 2.8)
hvatir fundu þat skatnar (ibid. 9.6)

If it is a fact that dróttkvætt lines follow this rule about the overall length of the line, counted in more, it makes them good evidence concerning the prosodic structure of the language. Just as we noted that if we were to find forms like tala regularly functioning as the last foot of a line we would suspect that a change had taken place either in the rules of the metre or in the

underlying language, we would suspect that a line with too few more according to our principle was composed after the lengthening of old short syllables had taken place. For example a line like:

hvatir fundu skatnar

which is exactly like the next line above except that one short monosyllable is removed so that the line only has eight more according to the old principle, would arouse our suspicion, since it does not fit unless the first syllable of hvatir is taken as long.

In search for evidence of the lengthening of old short vowels we will, then, have a look at dróttkvætt-poetry from different times in order to see whether we can find any signs of changes in the pattern we saw above.

To do this I have investigated the following material:

1. Glymdrápa by Þorbjörn hornklofi from about 900 (Jónsson 1912-1915 AI:22-24). (64 lines.)
2. Lausavísur (occasional stanzas) composed in the dróttkvætt-metre and ascribed to Egill Skallagrímsson in Egils saga (Jónsson 1912-1915 AI:48-59). These allegedly date from the 10th century, but the authenticity of some of them has been questioned. (340 lines.)
3. Erfidrápa Ólafs helga by Sigvatr Þórðarson, from about 1040 (Jónsson 1912-1915 AI:257-265). (206 lines.)

4. Hákonarflokkur by Sturla Þórðarson, probably composed in the years 1263-1264 (Jónsson 1912-1915 AII:124-127). (84 lines).
5. Pétresdrápa, an anonymous poem from the 14th century (Jónsson 1912-1915 AII: 500-508). (424 lines).
6. Guðmundardrápa by Einar Gilsson from the 14th century (Jónsson AII:397-404). (320 lines).

This makes a total of 1438 lines.

As we have said before, we will be looking for lines where syllables that are short according to the principle outlined in Section 0.1 seem to function metrically as if they were long. Obviously, we can't take every breach of the old rules as evidence of change; we will have to allow for a certain amount of inaccuracy in the use of the metres. It is difficult to judge how much deviation to allow for, since different poets will have differed in ability and meticulousness, and the same poet may have been more careless on some occasions than others. In order to get some idea about how much deviation to allow for before drawing conclusions about linguistic or metrical changes, the first three items on the list above were chosen from times when we have already assumed that the lengthening of short syllables had not taken place. In evaluating the evidence of the metrics we must of course also keep in mind that a change we detect in the correspondence between linguistic forms and metrical function can either be caused

by a change in the language or a change in the metrical rules. In the case of the dróttkvætt-metre, however, it seems rather unlikely that the metrical rules could be changed so as to allow short syllables to take the function previously held by long syllables only, because it seems that alternation between long and short syllables was one of the basic features of the metres, and a change in the rules of dróttkvætt giving short syllables the same value as long ones would seem to amount to nothing less than abandoning the basic principle on which the rhythm was based and creating a more or less new metre based on new rhythmic rules.

Let us start by looking at the older poetry. In the three first items on the list, I found altogether 26 lines where the number of moræ counted in the way we described above is too low and where a historically short syllable would have to be taken as long in order to make the line conform to the pattern we have described above. The total number of lines considered was 610, so that the percentage of lines potentially indicating a lengthening of short syllables is 4.3. In my investigation of Hákonarflokkur by Sturla Þórðarson from shortly after the middle of the 13th century I found no lines which could be interpreted as showing signs of a lengthening of old short syllables. The total number of lines was 84.

In the 14th century poetry the results were the following: The total number of lines considered was

744. Of these 744 lines, 21 had too few moræ, according to our system, and could therefore conceivably be taken to indicate that a lengthening had taken place. The percentage is 2.8, lower than for the poetry from the 10th and 11th centuries.

Even though these statistics seem unequivocally to indicate that no change had taken place in the time that passed from the 11th to the 14th century and need perhaps no further comment, I would like to point out one interesting fact. This is, that almost all the deviant examples from both before and after 1200 involve monosyllabic words. That is, by assuming length for old short monosyllables, the right number of moræ comes out. Typical examples are:

ek bar sauð af nauðum (Lausavísa by Egill
Skallagrímsson:33.8)

where the short stem ek 'I' would seem to have to be taken as long to get a regular — — — — —
scansion, and:

barðisk vel sá's varði (Egill 8.3)

where vel 'well' will have to be taken as long. (It is assumed that the form sá's (contracted from sá es) is metrically unstressed and therefore cannot carry an ictus.)⁷⁾ This might perhaps lead one to suggest that the monosyllables had undergone a lengthening already in the 10th century (if the poetry is authentic). But if we look closer at the data, this becomes less plausible. In the

examples from the oldest poetry (Glymdrápa, Egill's Lausavísur and Sigvatr's Erfidrápa) these short syllables are always followed by words beginning in consonants, and it has been suggested (cf. Craigie 1900: 370) that such short monosyllables could carry a full ictus if they were followed by consonants. In the cases above, the sequences ek har and vel sá's could perhaps have been used to form regular feet by taking them as one compound, so to speak, using the initial consonant of the following word to take the place of a second consonant following the stressed vowel.

In the younger poetry, Guðmundardrápa and Fétrsdrápa, the situation is not exactly the same, since out of the 21 examples apparently involving short syllables carrying a full metrical ictus, only 14 are of the same sort as the examples cited from Egill's poetry, two involve bisyllabics, and five involve monosyllabics followed by vowels. The examples involving monosyllables followed by vowels are all from Fétrsdrápa. They are:

faðir gat son án sáði (2.3)

brjóst ok har inn löstu (4.3)

sínum vin ok tínr (26.6)

reistr af kör ok Kristum. (54.3)

kemr (or kemur) í stað at styðja (48.4)

It seems that in these lines the underlined words would have to be taken as long in order to get a regular scansion.

The two lines where short syllables in bisyllabic words apparently function metrically as long are also from Fétrsdrápa:

ótt til grafar dróttins (45.8)

sonr gleði var sveinum (19.7)

In the first line, the word grafar would seem to have to be taken to have a long syllable if the line were to get a regular scansion, and in the second line, either the first word, sonr will have to be taken as bisyllabic: sonur (assuming that the epenthetic vowel which developed between a final r and a consonant had already appeared) and with a long first syllable or gleði will have to be taken as having a long first syllable.

In general, it seems that the examples we have found where the rules that we have assumed determined the metrical function of long and short stressed syllables are broken are too few for us to be able to assign any significance to them. It may be tempting to suggest that the fact that short monosyllables seem to break the rules five times and short bisyllabics seem to break the rules twice in Fétrsdrápa indicates that there is a change under way and perhaps that the monosyllables tended to lengthen earlier, but we would want to get clearer statistical evidence before drawing any such conclusions, and we must remember that there were no examples to be found in Guómundardrápa where the old rules seemed to be broken (granted that a short monosyllable was allowed to carry the ictus when

followed by a word beginning with a consonant), and the vast majority of occurrences of old short syllables conform to the old rules.

2.2.1.2 Rímur

Since very little skaldic poetry exists from after the period to which Guðmundardrápa and Fétrsdrápa belong (roughly the 14th century) we will have to look for other sources of evidence after this time. We have just seen that the 14th century skaldic poetry we have looked at does not show clear signs of a change having taken place in the prosodic system of the underlying language, so any evidence as to when the lengthening took place is still to come.

Fortunately for us, another genre of poetry, which arose (probably) in the 14th century, can be used as evidence about the prosody of the language. This poetry was the rímur 'rhymes', a type of literature which was to flourish in Iceland all the way down to the present century. The rímur are usually considered to be of mixed origin (Þórólfsson 1934:35-51 and Craigie 1952, I:XII-XVIII). The metres, which are quite varied, derive many of their peculiarities, such as alternative end-rhyme and the number of lines to each verse (most normally four), from foreign folk ballads, but other features, such as the use of alliteration, are of native Icelandic origin. Furthermore, the poetic language is a direct inheritance from the skaldic poetry. The most common

and basic, and probably the oldest of the rímur-metres is the ferskeytt (meaning etymologically 'four cornered') -metre. A typical example is the first stanza of Ólafs ríma Haraldssonar, which is among the oldest specimens of this genre, dated about 1350-1370:

Ólafr kóngur örr ok fríör
átti Noregi at ráða
gramr var ei við bragna blíör

borinn til sigrs ok náða (For reference, see below.)

As can be seen, the alliteration follows the same principles as in the dróttkvætt-metre. There are two stuólar in the first line of every pair, and one höfuðstafr in the beginning of the second line: Ólafr/örr/átti and bragna/blíör/borinn. The main innovation, compared to dróttkvætt and other older Icelandic poetry, is the number of lines and the nature of the rhyme, alternative end-rhyme.

What we are particularly interested in about this metre is its rhythmic structure. A look at the stanza cited above will give us some idea about the rhythm of the ferskeytt-metre. In the first line we see that there is a regular alternation between long stressed syllables and short (metrically) unstressed ones. We can scan this line on the same principles as we used for the dróttkvætt lines:

Ólafr kóngur örr ok fríör

This line looks very much like a dróttkvætt-line except that one (in this case long) syllable has been added at

the end. The second line, however, is of the same length as a dróttkvætt line, counted in moræ:

átti Noregi at ráða

(We assume that there is a liaison between Noregi and at.) The third line scans exactly like the first one:

gramr var ei við bragna blíðr

and the last line scans like this:

borinn til sigrs ok náða

It is worth noting that in the stanza quoted above, two short syllables can equal one long one. If we, for example, analyze the odd-numbered lines into three feet, each with three moræ, we see that these three moræ can either be covered by one long stressed syllable followed by a short one, giving a trochee as in átti or by three short and/or unstressed ones, like borinn til. What is most important for us is that in Ólafs ríma and the other rímur from the oldest period, it seems to be a rule that if for example a line starts with a bisyllabic word with a short first syllable according to the old rules, a third syllable follows before the start of the next foot. This means that a short bisyllabic word like taka could not form a trochee, and a third syllable had to be added. This is why in lines like the following:

Hildings taka þá helgan dóm (Ólafs ríma 62.1)

we get an unstressed þá 'then' after taka in order to form a full foot. The form taka alone is not sufficient, as are Hildings and helgan. The temporal adverb is not strictly necessary for the meaning of the sentence, and

if taka could have formed a regular trochee, the poet could (if the metrical rules had allowed it) have had the line as follows:

Hildingr taka helgan dóm

which is a perfectly normal line after the first syllable of taka has been lengthened.

Obviously, lines like the one we just constructed are the sort of things we will be looking for as evidence that the lengthening of old short syllables had taken place. If we, for example, find a poem where it seems to be regular that forms like bera or a sequence like bar ek function as trochees, that is, without a third syllable following within the same foot, we will be tempted to assume that a lengthening of short stressed syllables had taken place when that poem was composed. To look for evidence of this sort I have studied the following rimur-material from the 14th to the 16th century:

1. Ólafs ríma Haraldssonar by Einar Gilsson (the author of Guðmundardrápa, cf. above) from about 1350-1370. (Rímnaefn I:1-9). (260 lines.)
2. Skáld-Helga rímur, from about 1400 (Rímnaefn I:105-165). The first 40 stanzas were investigated, giving a total of 160 lines.
3. Skjóðaríma, from about 1400-1450 (Rímnaefn I: 10-42). 100 stanzas were investigated, giving 400 lines.
4. Króks-Refs rímur, from the latter half of the

15th century. (Published in Króka-Refs saga og Króka-Refs rímur, 1883.) The first 102 stanzas were investigated, the total number of lines being 408.

5. Ólafs rímur Haraldssonar, composed after Rauðúlfs báttur, from about 1550 (Rimnasafn I:215-221). (Referred to below as Ólafs rímur B.)

All the 126 existing ferskeytt lines were investigated.

6. Pontus rímur, by Magnús Jónsson próði, composed in the years 1564-1566 (Pontus rímur, 1961).

696 lines were investigated.

Before we look at the evidence given by these data, we will make a few further remarks about the rhythm of the rímur.

It seems that the rhythmical rules of the ferskeytt-metre were not as strictly based on length as the rules of the dróttkvætt-metre seem to have been. It is quite common in ferskeytt for a stressed long syllable forming an ictus to be followed by two (or even more) unstressed syllables, so that instead of trochees we get dactyls.

Examples of this are:

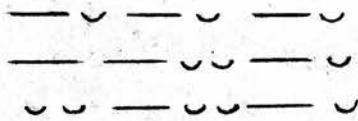
— — — — —
Fimm hefir kóngur kristnat lönd (Ólafs ríma Haraldssonar 4.1)

and:

— — — — —
grams fyrir merkit væna (Ibid. 38.2)

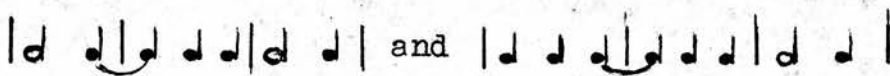
In both these lines, the first syllable is long, and consequently, only one syllable has to follow it in order

to form a full foot; but in both cases two short syllables follow. Since this sort of thing is quite common, it must mean that, unlike in the dróttkvætt-metre, the number of *more* in each line was not fixed. This is a quite important difference from the metrical point of view. Our hypothesis concerning the dróttkvætt-metre was that its rhythm was mainly based on length variation. We could get long and short syllables distributed over the line in remarkably many ways; we could get lines like:



etc.

In this type of rhythm, as we have seen, it is perhaps not appropriate to speak in terms of feet, but rather in terms of bars, as in music. Two bars can be linked together, one final note in the first bar being tied over the bar-line to a note in the second bar. The two latter types of lines can be represented as



respectively. In the rhythm of rimur, on the other hand, stress seems to have played a more important role. The fact that two metrically unstressed (usually short) syllables could quite regularly follow the stressed long one without disrupting the rhythm shows that it was not crucially based on the number of syllables or the length of the line measured in *more*, but rather that the basic,

or most central unit in the ferskeytt-metre was probably the foot with its ictus (Hebung) (which could be formed by one long stressed syllable or two short ones) and its drop (Senkung) which could be one or two (or perhaps even more) unstressed syllables. The fact that the ferskeytt-metre seems to have been basically different from the dróttkvætt-metre does not perhaps matter a great deal to us; but it is worth keeping in mind when we consider and evaluate the evidence of the ferskeytt-metre, that only a relatively minor change was needed in the rules to allow short stressed syllables to function as full ictuses. All that had to be done was to remove from the metrical rules a restriction forbidding a short stressed syllable to carry the ictus by itself. A change allowing for the same thing in dróttkvætt would have been much more drastic, since it seems that the whole rhythmic structure of the metre was based on length alternations. It seems, therefore, that when a change is detected in the relation between metre and language as far as prosodic length is concerned, such a change in ferskeytt has a greater chance of being caused by a change in the metrical rules than a similar change in dróttkvætt.

If we turn to the data we looked at, it seems clear that the same rule concerning the metrical value of short monosyllables is valid as in dróttkvætt; a short monosyllable could carry the ictus if it was followed by a word beginning in a consonant. Examples of this abound in

the oldest, as well as in the younger rímur. This means that lines like the following:

að þinn leiður lymsku drengr (Króka-Refs rímur I,68.3)
þat var mest af þryði (Ólafs ríma Haraldssonar 31.2)
innan hol sem kista (Skíðaríma 14.2)

must be ignored as evidence concerning the lengthening of short syllables and were probably quite regular before the quantity shift.

What we will be looking for, then, are examples where either sequences like $VC^1##VC_0$ or VC^1VC_0 function as whole feet. Assuming that the metre was unchanged, such examples can be taken as signs that the originally short syllables had lengthened. We will be interested to know when the short syllables lengthened and whether they lengthened at different times in monosyllables and polysyllabics.

In Ólafs ríma Haraldssonar only one example was found where it is possible that a short syllable would have to be taken as carrying an ictus in order to get a regular scansion:

þér innit framar hóti (32.2)

Here, it would seem most natural to assume that the pronoun þér is an unstressed upbeat, which is quite common in rímur, and if this is so, the three remaining bisyllabics will each have to form a foot by itself, including framar which has a short first syllable. The reason why þér, which has a historically long syllable and could therefore carry an ictus if stressed, is probably to be

taken as an upbeat and ignored in the scansion of the rest of the line, is that innit carries the alliterative höfuðstafr, which matches two word-initial vowels in the preceding line, and it seems to be a rule that such höfuðstafir can only be preceded by unstressed upbeats. It seems, then, that framar breaks the rule that an ictus can only be carried by a long stressed syllable or jointly by two short ones. One example in 260 lines, however, seems far too little to be of any significance.

In the 160 lines from Skáld-Helga rímur (from about 1400), no examples were found where it is certain that the old rules were broken.

In the next item on the list, the first 400 lines of Skíðaríma, from the first half of the fifteenth century, seven exceptions to the old rules were found. Three of these involve old short monosyllables appearing in front of words beginning in vowels:

bað er hvorki skrum né skjál (7.3)

Ofrligt er um örleik þinn (26.3)

hann Leifi kvað ei liggja á (37.3)

In these three lines, the words bað, er, and kvað, respectively will have to be taken as carrying the ictus in order for a natural scansion to be obtained. The other four examples involve bisyllabics:

úr máta stór og mikil er (13.3)

Ísungs synir útar í frá (81.3)

eru margir meiri en þú (95.3)

Á Íslandi eru margir menn (96.1)

Here, the first syllables of the words mikil, synir, and eru (twice) will have to carry the ictus if a natural scansion is to be obtained. Out of 400 lines this does not seem to be impressive statistics, and undoubtedly these examples will have to be dismissed as not significant.

To test this, we may like to look at the data from a slightly different angle. In a special survey of the first 160 lines of Skíðaríma, I counted the overall occurrence of old short first syllables of polysyllabics in metrically stressed position, that is, within the ictus. Short first syllables of polysyllabics occurred 25 times in this position, and in 24 of these instances the shortness of the first stressed syllable was compensated for by the presence of an extra syllable in the same foot, that is, a bisyllabic like taka was followed by an unstressed word like það in order to fill the foot. Only once in these 25 instances was there no such compensatory syllable.

For the monosyllables, the comparable figures were not quite as convincing, and I took a bigger sample of 400 lines. In these 400 lines I found 59 instances of short monosyllables occurring under metrical stress in the middle of lines (we will later come to the behavior of old short monosyllables at the end of lines). Three of these occurrences are the ones mentioned above where a word beginning in a vowel follows, and the rules we have set up seem to be broken. Looked at in this way

they seem to be harmless for our model. But a closer look at the data arouses suspicion. We note that in 41 instances short monosyllables are followed by a word beginning with a consonant and only followed by one syllable before the beginning of the next foot, as in the line cited above:

innan hol sem kista (14.2)

When these 41 instances and the three we have already labelled as breaking the rules are added together, we see that 44 times out of 59 the shortness of a monosyllable is not compensated for by the presence of a third syllable within the same foot, and only 15 times do short monosyllables occur as the first syllables of trisyllabic ictuses. If we were to say that the instances where monosyllables carry an ictus when followed by a word beginning in a consonant are to be taken as breaches of metric rules, we would come up with a situation where old short monosyllables break the old rules in 44 instances out of 59. In that case, one might wonder whether the old monosyllables had already become long and the instances where they occur in trisyllabic feet are of the same sort as in:

Fimm hefir kóngur kristnað lönd (Ólafs ríma Har-
aldssonar 4.1)

where a long stressed syllable is followed by two short syllables within the same foot. To this hypothesis one must first make the (relatively weak) objection that already in the oldest dróttkvætt-poetry, it is quite

common for short monosyllables to function as if they were long, when followed by a word beginning in a consonant, whereas it seems to be an exception if they do so when followed by a word which begins in a vowel. This would be difficult to explain if this hypothesis were correct; it seems that at the earlier stages there was some rule which allowed short monosyllables to carry the ictus when followed by a word beginning with a consonant and it seems unlikely that this derived from the fact that short monosyllables were longer when preceding a word beginning in a consonant. The discrimination against old short monosyllables when followed by words beginning with vowels seems to indicate that they were linguistically different from long ones. Another way of trying to see whether this hypothesis is likely to be true is to compare the function of old short monosyllables in Skíóaríma to their function in some other poem which dates from a later time, when we have reason to believe that a change had taken place in the metrical function of old short stressed syllables. Pontusrímur (see below) are considered by Björn K. Þórólfsson (1934:292-293) to be composed after the quantity shift took place, and we will consider a sample of 140 lines from this poem. The comparable facts of Pontusrímur show a different statistical pattern. In the 140 lines examined for this purpose, old short monosyllables occurred 21 times metrically stressed in the middle of lines. In 20 instances their shortness was

not compensated for by the presence of a third short syllable and would therefore have to be taken as carrying the ictus by themselves. The ratio for Pontusrímur is, then 20/1, whereas the comparable ratio for Skíðaríma is 44/15. Similarly, in Pontusrímur, the number of times an old short syllable carrying an ictus was followed by a consonant was 8, whereas it was 12 times followed by a word beginning with a vowel. The comparable ratio for Skíðaríma is 41 times followed by a word beginning with a consonant and 3 times followed by a word beginning in a vowel. The difference between 41/3 and 8/12 is so substantial that it seems reasonable to assume that a change had taken place in the interval that passed between the time when Skíðaríma was composed (in the first half of the 15th century) and the time when Pontusrímur were composed (around the middle of the 16th century). Since, in Skíðaríma, an old short monosyllable, seemingly carrying the ictus, was followed only three times by a word beginning with a vowel, it seems that there was a restriction against it which later was relaxed.

In Króka-Refs rímur, which are considered to be from the latter half of the 15th century, two examples were found where an old short monosyllable seemed to carry the ictus when followed by a word beginning with a vowel, and four examples, where an old short first syllable of a polysyllabic seemed to fill an ictus, showed up. The total number of lines examined was 408,

and when this is compared to the three monosyllables and four bisyllabics that seemed to break the rules in the 400 lines of Skíðaríma, we will have to assume that no change had taken place.

When we turn to the next item on the list, Ólafs rímur B, which are considered to be from about 1550, we seem to detect some change. Only 126 ferskeytt-lines are preserved from these rímur. They were all investigated with the results that four examples were found of old short monosyllables seemingly carrying an ictus when followed by words beginning with vowels, and 12 examples were found where a first short syllable in a polysyllabic had to assume the status of a full ictus before a natural scansion could be obtained. When we consider that the number of lines is only 126, approximately one third of the material from Skíðaríma and Króka-Refs rímur, this seems to indicate a change; we will have to multiply the figures from Ólafs rímur B by three to compare them with those of Skíðaríma and Króka-Refs rímur, making the figure for monosyllables 12, and the figure for polysyllabics 36. The comparable figures for Skíðaríma and Króka-Refs rímur were 3 and 4, and 2 and 4 respectively. These figures indicate, though perhaps not conclusively because of the poverty of the data, that a change was taking place or had already taken place about 1550.

It will have been noticed that the number of exceptions to the old rules involving polysyllabics is

considerably higher than the number of exceptions involving monosyllables. This might lead someone to suggest that a lengthening had hit polysyllabics earlier and more regularly than monosyllables. But it is doubtful that any such conclusions can be drawn from these data, since, in general, the incidence of polysyllabic words is greater. (In a randomly chosen prose text of 23 lines, I counted 146 polysyllabic words, but only 91 monosyllabic ones, and in 14 stanzas from Ólafs ríma Haraldssonar, I counted 145 polysyllabic words against 111 monosyllables. In both cases, unstressed words like prepositions and conjunctions, which are mostly monosyllables, were included.) These figures, then, don't allow us to assume that there was any time interval between the changes of the metrical function of old short syllables, according to whether they occurred in mono- or polysyllables.

We have already had a look at the evidence of Pontusrímur. In our main sample of 696 lines (the 1st and 11th ríma) 22 examples were found of old short monosyllables carrying an ictus, and 41 of historically short first syllables of polysyllabics. Although these figures are perhaps not as striking as those we have already seen from Ólafs rímur B, they undoubtedly show that the old rules no longer prevailed. If we look first at the polysyllabics, we see that in 696 lines we get 41 examples of short syllables functioning as long. That means that there is one breach of the old rules in

approximately 17 lines (41:696); in Ólafs rímur B there is one breach of the old rules per approximately 10 lines (12:126). Both of these figures show a significantly higher incidence of breaches of the old rules than the figures for Skíðaríma and Króka-Refs rímur, where there is one breach for approximately every 100 lines (4:400 and 4:408 respectively). It seems, then, relatively safe to assume that the old short polysyllabics had a different function in Ólafs rímur B and Pontusrímur from that in the 15th century poems, Skíðaríma and Króka-Refs rímur. The same can be said about old short monosyllables. Here we have, in Pontusrímur, one breach for approximately every 31 lines (22:696), and a similar ratio shows up in Ólafs rímur B (4:126), but in Skíðaríma there is one breach for approximately 133 lines (3:400), and one for approximately 200 lines in Króka-Refs rímur (2:408).

2.2.1.3 The last feet of untruncated lines

From our investigation of the evidence of rímur-metrics so far, it seems clear that around the middle of the 16th century there occurred a change in the metrical function of old short syllables, so that the restrictions against their occurrence in place of a monosyllabic ictus in a trochee were relaxed. Forms like tala, vera, bar ek, etc. are now allowed to form a full foot, something they could not do before. This was investigated earlier by Þórólfsson (1934:291-294 and 1929a)

with similar results. Þórólfsson took this as evidence that the old short syllables lengthened around the middle of the 16th century (at least in the dialects represented by the poems in question), in other words that the change was a linguistic one, although one cannot of course exclude the possibility that what took place was simply a change in the rules of the metre. We have already mentioned that a relatively small change was needed in the metrical rules for rimur to allow for this change. The old condition for a syllable to be able to carry the ictus by itself was that it was stressed and not short. (There were even exceptions to the length condition, in that a short monosyllable could carry the ictus if followed by a word beginning with a consonant.) All that had to be done to allow for the change that we have detected was for the condition concerning length to be removed. If this was the case, there is no need to assume that the change that took place about 1550 was the linguistic lengthening of short vowels. Indeed, it may seem just as likely that the change was simply a change in the metrical rules (cf. Kjartansson 1971). If we look at the last feet of untruncated lines (lines ending in bisyllabics, i.e. trochees), we see that the old restriction concerning the length of the stressed syllable in this metrical position seems to have prevailed much longer. Stefán Karlsson (1964:10-11, fn. 12) notes that, in the whole of Þortusrímur by Magnús Jónsson þrúði, old short bisyllabics occur only 10 (or 9) times as the last feet of untruncated lines. The

total number of untruncated lines is 1213, which gives us about one breach in every 120 lines. This seems a considerable consistency in distinguishing between old short and long stressed syllables in bisyllabics and is difficult to explain if no linguistic difference is assumed to have prevailed between these two types of syllables. Karlsson also shows that in a rímur-poem from the first half of the 17th century (composed in 1643, according to the only existing manuscript) Egils rímur Skallagrímssonar by Jón Guðmundsson, a similar rule is maintained. Only in about 1% of lines ending in bisyllabic forms is the first syllable short according to the old rules. In these rímur, as well as Pontus-rímur, there seems to be no restriction against old short syllables carrying the ictus by themselves when they occur in a non-final position in the line, but the fact that old short bisyllabics hardly occur as last feet of untruncated lines in these poems is difficult to explain if no distinction was made in the metrical rules between old long bisyllabics (hestur, óska) and old short bisyllabics (taka). As long as a distinction was made between the metrical function of these two types of syllables, one must, it seems, assume that some relatively simple linguistic feature existed in terms of which the metrical rules were defined. It makes no difference which aspect of the metrical rules is based on this distinction; as long as a distinction is made somewhere in the metrical rules it must be assumed that

some linguistic difference existed. The 16th century change may then have been simply a change in the metrical rules as far as Pontusrímur are concerned, since the author, Magnús Jónsson prúði, seems to have little or no trouble distinguishing between the two types of syllables when it comes to deciding which words are eligible as the last feet of untruncated lines.

We might then want to say that the lengthening of old short syllables did not take place until the 17th century at the earliest. But if we look closer at the data investigated by Þórólfsson (1929a), this becomes less attractive. We have noted that a relatively minor change was needed in the metrical rules for ferskeytt (and other related metres) for old short stressed syllables to get the same metrical value as old long ones, but I have also suggested that such a change in the metrical rules for dróttkvætt would have been more drastic. It is interesting that the poet Hallur Ögmundsson (according to Þórólfsson, one of the first to show signs of a change in the function of old short syllables in his poetry) breaks the old rules of the hrynhentmetre, which was a version of the dróttkvætt, developed in the 11th century (Þórólfsson 1929a:40-44). Not only are there several examples where it seems that old short syllables in non-final position in the metrical line will have to be taken as long in order for a natural scansion to be obtained, but more strikingly, there are quite a few instances of old short bisyllabics functioning as

last feet of the lines, which, as we have seen, seems in general to be the most conservative position as regards distinction of length and shortness of stressed syllables. Þórólfsson (1929a:51) also notes that Einar Sigurðsson (1538-1626) breaks the old rules in dróttkvætt, (and another old metre, fornyrðislag). If the 16th century change was a metrical one, it must, then, have been a change not only in the rules for the rímur-metres, but also in the rules for the old dróttkvætt and hrynhent (as far as they were still in use). It may seem a strange coincidence that changes should be made in all these metres at the same time. Why should the poetic conventions be changed so drastically in the 16th century?

We have here the following puzzle: A partial change takes place concerning the metrical function of old short syllables. The question is: was it a linguistic change or merely a change of metrical rules? If it was a linguistic change, then why did the two types of old stressed syllables (long and short) still have different metrical value in some positions in some metres? On the other hand, if the change was metrical, then (apart from the question of why it took place) it may seem a bit of a remarkable coincidence that more or less all metres were hit by the same sort of change at the same time.

It is interesting to note that, whereas the above mentioned Einar Sigurðsson breaks the old rules concerning the metrical function of old short syllable in

dróttkvætt and other old metres, he maintains them in the last bisyllabic foot in rímur and other younger types of poetry. These younger metres have in common the fact they were intended for singing (or chanting). The fact that the old distinction was still maintained only in chantable metres, may partly resolve our paradox. If the 16th century change was a metrical change partly caused by a lengthening of old short syllables, it is possible that it was still not permitted that these old short syllables be stretched in singing or chanting, as for example in the rímur, assuming that the last feet of untruncated lines typically occurred on long notes in the appropriate tunes. We don't, in other words, have to assume that the fact that these old short syllables did not occur in the last feet of untruncated lines which were intended for singing or chanting, necessarily means that the old short syllables had not lengthened at all. It is conceivable that for some reason they were not fit for being stretched in singing. One could perhaps imagine that there still was some difference between the vowels of tala and lúka as far as 'stretchability' was concerned although the difference was not so great as to forbid the old short syllables from functioning as an ictus in the middle of lines. It is possible that only when it came to stretching the syllable in singing in final position in a line that a difference was detected between old short and old long syllables. It is a characteristic of many of

the tunes to which the rimur were chanted, that they have long notes in the position of the last ictus of even numbered lines (which are always untruncated in simple ferskeytt and many other rimur-metres). A typical tune, transcribed in a 4/4 rhythm by Þorsteinsson (1906-1909) is the following:

Und - ir steini á há - um hól,
hlýt ég reyn - a kul - da,
var - la hreina séð fæ sól,
svið - ur mein - ið duld - a

(Þorsteinsson 1906-1909:871)

And in a 3/4 rhythm a typical tune is the following:

Vatns - dæl - ing - ar veita' ó - spart
veg - far - end - um bein - a,
elsk - a heið - ur, hefð og skart og
hrund - ir eð - al stein - a.

(Þorsteinsson, op.cit:873)

Both of these tunes show that the last feet of untruncated lines are particularly long in singing, and it is to be expected that if there were any restrictions on the occurrence of old short syllables, these should be the positions where they are most consistently forbidden.

I will not try to give a conclusive solution to our puzzle here, since a much more detailed investigation of the data is needed than can be done in the context of this study, but merely state what I think is likely to have been the course of events and comment on possible ways of trying to solve the problem of the dating of the lengthening of old short stressed syllables.

I find it likely that the change that can be seen in the 16th century poetry is a consequence of a gradual lengthening of old short stressed syllables, most typically by a lengthening of the vowel, but possibly also by a lengthening of consonants in some cases. This lengthening was, however, not so great as to make it possible to use the old short syllables to carry the long notes that characterized the last ictuses of untruncated lines in metres that were intended for singing, such as rimur and some other metres, mostly of foreign origin. In other words, my answer to the question whether the lengthening of old short vowels took place in the 16th century is yes, to a certain extent, but not completely. The old short syllables had lengthened, at least in certain dialects, but there could still be

detected a difference between them and the old long syllables. I am suggesting, then, that the lengthening of old short syllables and the change in their metrical function was a gradual process.

The idea of a gradual lengthening of old short syllables fits in well with the fact that the disappearance of the 'discrimination' against short syllables in the last feet of untruncated lines seems to be gradual. As we have already seen, there was one short bisyllabic last foot for approximately 120 lines in Pontusrímur. A similar ratio (1/125.) holds for Egils rímur (1643) (Karlsson 1964:22). In Olgeirs rímur danska by Guðmundur Bergþórsson (from 1680) the ratio is 1/47. In Brávallarímur by Árni Böðvarsson (from 1760) the ratio has reached 1/12, and in Númarímur by Sigurður Breiðfjörð (1833), it is 1/5.6.⁸⁾ These figures show clearly that it gradually becomes more natural to have old short syllables functioning as ictuses of the last feet of untruncated lines, until there is no restriction against it.

It is perhaps worth noting that the last foot of untruncated lines can only be used as evidence about bisyllabics; we have no evidence concerning monosyllables. For all we know, they may have lengthened earlier than the bisyllabics, but we have no evidence for this, one way or the other.

Another point worth making, concerning the progress of the lengthening of old short syllables, is that it is quite likely that it progressed at different speeds in

different dialects. Þórólfsson (1929a:79) suggested that the lengthening was an innovation that spread from the West, Hallur Ógmundsson being from the West. Karlsson (1964:23) points out that this does not fit well with the fact that Jón Guðmundsson í Rauðseyjum, the author of Egils rímur, also from the West; still distinguishes between long and short syllables in 1643. It is evidently necessary to make a thorough investigation of poetry from the 16th and 17th century before any conclusions can be drawn as to the geographical progression of lengthening of old short syllables.

Before concluding this subsection let me make a few comments concerning the relation between the lengthening of vowels in monosyllables and polysyllabics. At the beginning of this section we mentioned two alternative models, what we called the 'overnight hypothesis' and the 'conspiracy hypothesis.' These alternatives are both based on the assumption that changes are abrupt (cf. footnote no. 5). We set up a choice between lengthening of vowels in all short syllables at the same time (overnight) or in two steps (conspiracy). One might now say that the distinction as far as vowel lengthening is concerned between 'overnight' and 'conspiracy' has evaporated, since we have abandoned the idea of abrupt lengthening. Since we have started talking about a gradual lengthening of short syllables, it may seem of little value to distinguish between two subparts of a change that takes about two centuries to be completed. One might say that

it would be of no significance, though vowels of monosyllables tended to lengthen earlier than vowels of bisyllabics, if the whole change was a gradual move from two to one length of stressed syllables.

When we talked about conspiracy or a single change, we talked as if it was quite clear what is meant by a single change. "It is worth some consideration what conditions must be filled by something in order that it be called a 'single change'. It has become increasingly clear recently that linguistic innovations can take a long time to be carried out; the changes are typically gradual, as it has been put. We notice immediately that 'gradual' can be used to refer to linguistic changes in different ways, depending on which way one looks at it (cf. Wang 1969). There are at least three senses in which phonological changes can be gradual (and, logically at least, the opposite: abrupt). One can say that a phonological change is phonetically gradual, meaning that phonetic (allophonic) changes that can be said to underlie the structural changes take some time to develop. This change can proceed 'gradually' either in the speech of one speaker or in the speech of a community. Another dimension along which graduality can be measured is the inter-individual dimension; one can grade the speed by which some innovation spreads from one individual to another, e.g. according to class or geographical location. A third dimension is the vocabulary of the language. It is possible that some changes are more regular in some

parts of the vocabulary than others and that it takes some time for the change to reach all forms of the language. Indeed, it is quite common for some special parts of the vocabulary never to be affected by the change; they become exceptions to synchronic rules (cf. Ralph 1975:132-162). When we maintain that the lengthening of short vowels was a gradual change, we will do well to clarify in which of the above mentioned senses it was gradual. The answer is that it most likely was so in all three senses. If we assume that the lengthening was connected with the stressedness of the syllables, it is very likely that it was phonetically gradual, even to the extent that individual speakers gradually lengthened their originally short stressed vowels as they got older. It is also quite likely that the innovation gradually spread from one geographic area to another, and perhaps also from one class to another. Thirdly, it is quite possible that the vowels lengthened earlier in one part of the vocabulary than in others, and the change was thus 'lexically gradual'. Looking at the nature of the lengthening, it seems to be such that it is very likely to have proceeded gradually. Most importantly, it did not produce any clashes in the system, since quite early, perhaps as early as in the 13th century, the distinctive function formerly carried by vowel-length began to be taken over by qualitative differences, so from that point of view, the length of segments was redundant, and speakers were free to vary it. There was thus no hurry to restore

order to the minor irregularity created by varying length of vowels. The phonetic nature of the change (the phonetic environments for the linguistic change, if you like) is also such that it can easily be conceived as progressing gradually. The only place where there is any reason to introduce abruptness into the whole business is the assumed structural innovation made by speakers when they start making generalizations about the length of stressed vowels according to environment; when we, ideally speaking, have one generation (one speaker) having a more or less regular distribution of length without having strict phonological rules for it, followed by another generation (another speaker) with an underlying phonological rule (or rules) accounting for it. (This is what H. Andersen 1973 calls 'abductive change.')

But of course the spread of this innovation can still be gradual on the inter-individual (dialectal) parameter.

If the change was thus gradual in almost every sense of the word, it may seem, as we said before, that it makes no sense to speak in terms of a distinction between a single change or a conspiracy as far as the lengthening in monosyllables and polysyllabics is concerned. There is, however, nothing wrong with talking about two gradual changes, one beginning earlier than the other, or having an earlier 'center of gravity' than the other, even though the second change may have started before the other one was completed.

Also, it seems to make sense to think of new generations making different phonological generalizations at different times. For example: 1. All stressed vowels are short in front of two or more consonants. 2. All stressed vowels are long before one (or no) consonant in monosyllables. 3. All stressed vowels are long before single consonants in polysyllabics. If these generalizations come up at different times, they can be called three different changes or at least three stages in a single change. This is essentially what our 'overnight' vs. 'conspiracy' distinction is based on. Regretfully we have not been able to find any evidence as to whether there was any difference in time between the beginnings or 'the centers of gravity' of the two lengthenings, but the question is quite clear. It is another question, different from the gradualness question, how much time to require between signs of the beginnings of potentially separate changes, and, perhaps more importantly, how much difference of nature, causes, etc. to demand for two events to be called separate changes. We will return to the question of conspiracy in Section 3.4.

2.3 The environment p, t, k, s + v, i, r.

We have seen (Chapter II, Section 3) that there is an exception to the length rule in Modern Icelandic in that in front of two-consonant sequences, of which the first member is p, t, k or s and the second v, i or r,

vowels are long, whereas the general rule is that vowels are short before two or more consonants. We suggested that this was due to the fact that the p, t, k, g + v, i, r-sequences are syllabified differently from other sequences of two or more consonants. We proposed that the forms vekja (with a long vowel [vɛ:c^ha]) and vakna (with a short vowel [vahkna]) should be syllabified vekja and vakna respectively, and that the difference in syllabification could either be explained as a consequence of a strength hierarchy in the consonants (v, i, r being 'weak' in that they have a tendency to stand near the vowel (nucleus) of the syllable they belong to) or by a syllabification rule based on a distinctive feature classification of the consonants (making v, i, r the only underlyingly voiced fricatives in Modern Icelandic). In this section we will look at this problem from the historical point of view.

Since forms like vekja 'wake up' (transitive), vökva 'to water', akri 'field', (dative), setja 'to put', þrisvar 'three times' etc. show, in Modern Icelandic, long reflexes of old short vowels, we must assume that a lengthening of vowels took place in front of p, t, k, g + v, i, r-sequences at some stage.

It is interesting to see what function these forms had in the older poetry, that is, whether their first 'syllable' functioned as long or short. It turns out that they seem to have functioned as long. For example in the poetry of Sighvatr Þórðarson (from the 11th

century, see Jónsson 1912-1915 AI:223-275) there are several examples of forms of this sort filling the last foot of the line in dróttkvætt, which as a rule is only filled by bisyllabic forms with long first syllables. (E.g.: frændsekiu styr vekja (Flokkur um Erling Skjálgsson 7.4; Jónsson 1912-1915 AI: 246) and Hét sá er fell á Fitjum (Bersögglisvísur 4.1; Jónsson 1912-1915 AI: 252).) In rímur-poetry from the 14th and 15th centuries the same seems to be the case. The vekja-type words have the same function as other words with two consonants following the vowel, that is, there are several examples where their first syllables carry the ictuses of trochees. It is the case, then, that the sequences $V + \{p/t/k/s\} + \{v/j/r\}$ have the same metrical status as other sequences of vowels plus more than one consonant.

Before going any further we will note that the fact that these forms function as long in the metre seems to suggest that the segments i, y (and r) had the same phonological status as other consonants. It seems that it is wrong to regard them (at least in this environment) as semivowels as has often been done in the case of i and y for the earliest stages of Icelandic (cf. Noreen 1923/1970:44-45). If i and y had been semivowels (or 'glides'), presumably definable as nonsyllabic vowels, they would most properly have been analyzed in forms like setja as the first components of rising diphthongs following the consonant, that is, setja would consist

phonologically of a sequence CVCVV or CVCSV (S= semi-vowel), something like /setia/, /ia/ being a rising diphthong. If it were the case that i and y were non-syllabic vowels in Old Icelandic (11th to 15th century) one would expect the first syllable of the vekia-type words to function metrically as a short one, having a stressed syllable with only one consonant following the vowel. But this is not the case. (I will return briefly to the problem of the phonological status of semi-vowels in Chapter V, Section 4.)

The metrical function of the vekia-type words, then, seems to group them with long-stem bisyllabics as far as prosodic structure is concerned. (There are no monosyllables ending in a y or a i following p, t, k or s, and words like akr 'field' with an r following a consonant word-finally became bisyllabic in the 13th or 14th century by the rise of an epenthetic vowel between r and the preceding consonant.) But this does not fit with the situation in Modern Icelandic, where the forms show reflexes of a vowel lengthening of the same type as in short-stem words. It seems that the vekia-type words changed groups, so to speak, having been long-stem bisyllabics in Old Icelandic, joining the old short-stem bisyllabics in the quantity shift. This we will have to try to explain. We would like to be able to distinguish the p, t, k, s, + y, i, r-sequences from other postvocalic consonant sequences and explain why vowels lengthened in those environments.

Before considering this more closely we should point out that a priori there is more than one way in which a lengthening of vowels in front of p, t, k, s + v, i, r could be related temporally to the lengthening of short vowels in other environments. It is possible that before the lengthening of short vowels occurred, a split took place in the set of post-vocalic consonant sequences, so that the p, t, k, s + v, i, r sequences came to provide a favorable environment for historical vowel-lengthening, but the others did not. It is not, however, necessary to assume this; it is equally possible that the split in the post-vocalic sequences took place after the historical lengthening of old short vowels had taken place. If this was the case, the change in the consonantal sequences simply created a new environment for a synchronically active lengthening rule. A third alternative is that the vowel-lengthening and the change that caused the split in the post-vocalic consonant sequences took place at more or less the same time. If the first of these alternative chronological orders of events is the right one, one could expect that before signs of the vowel lengthening showed up in the poetry a change could be detected in the metrical function of the vekja-type words. That is, if for example some sort of reorganization in the syllabic structure took place, making the syllabic structure of the vekja-words identical with that of taka, but different from that of e.g. kasta, this could conceivably show signs

in the poetry in that vekja etc. could not function as trochees any more. I have found no signs of this in the poetry. We have, then, no external evidence as to whether a change in the prosodic status of p, t, k, s + v, i, r sequences took place before or after or at the same time as the lengthening of short vowels.

In the absence of direct evidence, we seem to be allowed the privilege of speculating freely about what happened. If we stick to our hypothesis set forth in Chapter II that in Modern Icelandic stress determines syllabification which in turn defines the scope of the length rule and we assume that stress 'caused' the lengthening of old short vowels, we can speculate about the development in the vekja-type words and the relation of syllabification and stress to the quantity shift.

We may say that the fact that the vekja-type words seem to have the same function as regular long bisyllabics in the metrics before the quantity shift indicates that the concept 'long syllable in Old Icelandic', as we have used it in our discussion of metrical function, was something quite different from the Modern Icelandic (necessarily long) stressed syllable. Indeed, it could be argued that the term 'syllable' should not be applied to the Old Icelandic unit at all. In our discussion we have simply used the term (stressed) syllable to refer to the stressed vowel plus all the consonants that follow up to the next vowel. (Of course, the initial consonant would be included in the syllable.) As is clear from our

discussion in Chapter II (Section 3.) of the concept syllable, this use of the term is perhaps the least likely to have any appeal to linguists, since generally it seems that people tend to favor syllabification which maximizes open syllables (cf. e.g. Fulgram 1970). The motivation for our syllabification of Modern Icelandic - extending the domain of the stressed syllable as far back as possible (final maximalistic, cf. above) - was that by adopting such a syllabification we could define the domain of several phonological processes, among them the rule governing the length of stressed vowels. As far as we accept this motivation we accept the stressed syllable as a linguistically significant unit in the phonology of Modern Icelandic. As far as Old Icelandic is concerned, there are no signs to be found of the phonological processes we used in justifying the final-maximalistic stressed syllable as a unit in Modern Icelandic. On the contrary, the metrical behavior of the vekjá-type words can be interpreted as evidence against it. If we assume that the metrical rules of dróttkvætt and rímur operated in terms of linguistic entities, that is, they were of the general form:

Linguistic unit X has metrical function F,
and that the optimal relation between any linguistic entity and any metrical function was that X always had the same function, and conversely that F was always filled by the same entities, it would be natural for us to assume that the first four segments of setja as a sequence had

the same linguistic status (or non-status) as the first four segments of yakna. The sequence of these four segments is what we called a 'long stressed syllable in Old Icelandic'. But this 'syllable' must have been something different from the Modern Icelandic stressed syllable. Our Modern Icelandic syllable is defined by stress, that is, the stress pushes the syllable boundary as far back as it can according to the phonotactic restrictions on the order of non-nuclear (consonantal) segments. If the unit utilized by the metrical rules was a final-maximalistic syllable defined in terms of stress, like the Modern Icelandic one, that syllable would have to be defined so as not to distinguish between the vekja-type words and other words having two consonants following the stressed vowel. But it seems highly unlikely that a stressed syllable for Old Icelandic could be defined as final-maximalistic without a distinction being made between setja on the one hand and yakna and kasta on the other, since then the syllable boundary would have to be pushed back of the v, i, r of the vekja-type words. If a syllabification setja, akr*ði* etc. is unlikely to be valid for Modern Icelandic, it is even less likely to be valid for Old Icelandic, since this could only be justified by assuming that the stress pushed the syllabic boundary of the stressed syllable even further back in Old Icelandic than in Modern Icelandic. Indeed, as will be seen from what follows, it is my main thesis that the primary prosodic difference between Old and Modern Icelandic is that stress has come to play a more central role than it did before.

Although it seems unlikely that a linguistic unit of the same sort as the Modern Icelandic final-maximalistic stressed syllable existed in Old Icelandic, it is possible that some phonological unit of a different kind that could be called a 'stressed syllable' existed. It is, for example, conceivable that setja, vakna and hestur were all syllabified like this: vek#ja, vak#na, hes#tur. In that case the change from Old Icelandic to Modern Icelandic as far as syllabification is concerned consisted in pushing the syllable boundary back one consonant if not prohibited by phonotactic constraints of the type discussed in Section 3 of Chapter II. This may well have been the case as far as we are concerned, but before postulating this it seems reasonable to demand that some independent justification be given for it, for example by showing that some phonological processes are more easily explained if such a syllabification is adopted. (We work, in other words, on the principle proposed by J. Anderson (1969) that if some phonological rules or some other phenomena are more easily accounted for by adopting some particular syllabification or some principle for syllabification, that syllabification or syllabification principle should be adopted.) I have not found any purely linguistic phenomena which seem to demand such a phonological syllable for Old Icelandic. It is, however, interesting to note that if this syllabification is adopted, either at an abstract phonological level or on some lower phonetic level, one peculiarity in the metrical

behaviour of long and short 'syllables' can perhaps be explained. We have already mentioned that it seems to be an exceptionless rule that the last feet of dróttkvætt-lines as well as untruncated lines in rímur had to consist of a long-stem bisyllabic, like kenna, hestur, láta etc. The interesting thing is that when (particularly in rímur) the line was truncated (i.e. ended in a monosyllabic foot) there seems to have been no demand that that syllable was long according to the old rules. These lines could just as easily end in short forms like dag, tal as in long forms like gest, rik. (Cf. e.g. Ólafs ríma Haraldssonar 51.1: Þórir lagði í kóngsins kvið.) This seems to be strange from the point of view of what we have seen so far of the metrical function of long and short forms. We have seen that in the middle of lines old short stressed 'syllables', whether in mono- or disyllables, could not carry an ictus (with the exception that short monosyllables could carry the ictus if followed by a word beginning with a consonant). But the last syllables of truncated lines certainly seem to have carried metrical stress, for example since they usually rhymed with other line-final syllables, and the question is why short syllables should be able to carry an ictus in this position but not others. If a syllabification like the one just mentioned is adopted, an explanation of this can be proposed. If we suggest that the condition for something to be able to carry an ictus if the stressed vowel was short was that the syllable was closed, we could account for

the metrical behaviour of old short monosyllables and the vekja-type words. We can suggest that the vekja-type words were syllabified like this: vekkja, aktri, kesja etc., i.e. had closed first syllables, and by virtue of that fact could carry the ictus. The same could go for short monosyllables when they stood at the end of truncated lines; there was no way that kvið in front of silence could be syllabified other than kvið. Bisyllabic words standing at the end of (untruncated) lines, on the other hand, provided a following vowel so that if tala were to have occurred there, it would have given an open syllable: talla. When short monosyllables occurred in the middle of lines, their syllables were closed, we may suggest, when a word beginning with a consonant followed, but became open when a vowel followed. Thus, in this analysis, a sequence like hol sem would be syllabified holsem but a sequence like böl á would be syllabified bölla, when occurring in the flow of a poetic line. In this way böl á would have the same metrical function as bera, which would be syllabified bera. Needless to say, this is highly speculative, and various queries can be put forth.⁹⁾ In the first place, this only accounts directly for the metrical behaviour of words having a short vowel in their first syllable. It does not say anything about the syllabification and metrical function of forms like dæma, sækja or óska, which have long stressed vowels. But they cause no problems from the point of view of metrical function,

since, however they are syllabified, they will end up with at least a bimoric first syllable. The only forms that seem to be problematic for this sort of analysis are those with a vowel preceding a hiatus or a word-boundary, such as búa or bú, which function metrically as short in the oldest dróttkvætt poetry. But these forms are problematic anyway, since there seems to have been some doubt as to the phonetic length of vowels in this environment in Old Icelandic, and their phonological status was probably somewhat exceptional (see Benediktsson 1968).

Another sort of objection may be raised against my suggesting this explanation for the metrical behaviour of old long and short forms. It might be said that this explanation, if correct, would contradict what I have already said about the nature and causes of the quantity shift changes. I have suggested that the lengthening and shortening of Old Icelandic vowels was connected with a phonetic thing called the stressed syllable. If a phonetic unit of this sort is to be used to explain the lengthening of /a/ both in tak and taka, it will have to give the same environment in both cases, and this can best be done, it seems, by assuming that the syllabification was final-maximalistic, as in the ^{case of the} later phonological Modern Icelandic syllable. Also in forms like óska, the shortening of the vowel can only be explained if the second consonant following the vowel is taken into consideration; it seems to be crucial, because in the closed syllable, ós, no shortening takes place. This can be done

if the phonetic syllabification we use is final-maximalistic. I may seem to have led myself into a self-contradiction; I suggest one syllabification to explain the quantity changes and another to explain the metrical behaviour of linguistic forms of the same language at more or less the same time. This might seem to be sinfully opportunistic. A conceivable way out of this is to say that these two syllabifications are of a different nature. I have assumed throughout that there could be two different types of syllables: phonetic and phonological. A way of defining the difference between these two concepts is to say that for something to be a phonological syllable, it has to play a central role in the phonology of the language in that it, for example, figures in some phonological rules. In that respect the stressed syllable in Modern Icelandic would be a phonological unit. A phonetic syllable, on the other hand, can be looked on as a performance unit, only cropping up in the actual production of speech without playing a role in any structural phonological processes. The sort of thing we mentioned in trying to explain the nature of the changes in the duration of vowels would be a phenomenon of this sort. We might, then, want to say that the syllabification I just suggested to explain the metrical behaviour of old short monosyllables was not a phonetic but a phonological syllabification. In that way the two syllabifications could perhaps have coexisted before the quantity shift, and we could use the phonetic syllabification to explain the rise of the changes in vowel quantity, but the

phonological one to account for the metrical behaviour of old long and short forms. This is far from convincing, however. I have already mentioned that there is very little if any linguistic justification to be found for the 'final-minimalistic' syllabification in Old Icelandic. There is no evidence that the first three segments of tak on the one hand and taka on the other behaved differently in Old Icelandic, nor are there to be seen any signs of e.g. the t of kiöt 'meat' and jötni 'giant' (dative sg.) (syllabified according to the suggestion we are considering: kiöt, jötni) showing a common difference from the t of jötunn 'giant' (nominative sg.), allegedly syllabified jötunn. So, if this syllabification was phonological, it seems to have been pretty much without phonological consequences, and adopting such a phonological syllabification goes against our principles.

Is there, then, a way of using the 'open' (or 'final-minimalistic') syllabification to explain metrical behaviour and still sticking to the idea that the quantity shift had something to do with the existence of a phonetic final-maximalistic syllable? If we believe in the explanation of lengthening using the final-maximalistic syllable, we seem to be forced to admit that the 'open' syllabification was neither phonetic (because then we would not be able to explain the quantity changes in the way we have proposed) nor phonological (because we have no independent arguments for its existence). We may

seem to be left with no choice but to give it up. But there may still be a way out. It is logically possible that this syllabification only existed for metrical purposes, that is, it could have been imposed on the language only when it was used in poetry. In fact we have already suggested a kind of 'abuse' of the language in poetry. We suggested that the sequences ek bar and vel sá's, linguistically consisting of two different words, could be taken metrically as compounds, so that a sequence of a vowel plus two consonants was obtained, and that a sequence like böl á could be taken to form one 'metrical word', so that conditions arose for an 'open' syllabification. If we can believe that a sequence of two words could behave metrically as one, there seems to be little to prevent us from believing that a special 'metrical syllabification' could be imposed on the language when it was used in poetry. In this way we could maintain this syllabification in order to explain the metrical behaviour of some linguistic forms without having to assume that it had any linguistic consequences as an abstract phonological entity or was identical to 'every-day' production units and thus save the idea that the quantity changes were connected with a purely phonetic syllabification of a different sort. This 'metrical syllabification' could be a division of text into 'poetic production units'. This is of course still opportunism with respect to syllabification, but at least it is not self-contradictory.

We may leave it an open question what sort of phonological syllabification should be postulated for Old Icelandic. Indeed, it could have been something similar to the metrical syllabification we have been discussing, or something still different. There may even be no reason to assume that any phonological syllabification prevailed in Old Icelandic if no phonological phenomena need to be explained in terms of it. (This is rather unlikely, however, since it is probable that phonotactic rules could most easily be stated in terms of syllabification.)

It is perhaps worth emphasizing at this point that although I have suggested that there was perhaps a difference between the actual, let us say, 'every-day' syllabification and a 'metrical syllabification', that does not mean that the arguments I have put forth about the language, based on metrical function, are invalid. I am by no means suggesting that the language could be used in most any way in the poetry. The 'metrical syllabification' must, of course, have been defined in terms of the underlying language; it was, according to my assumptions, basically a division of the text into special production units which could fill the functions demanded by the metre. Just as the 'every-day production units' were originally based on the underlying linguistic forms, the 'poetic production units' were based on them too. It is therefore perfectly legitimate to use metrical function to make inferences about the phono-

logical structure of the language. It is a fact that tala had a different metrical status from that of dama, and it is reasonable to ascribe that difference to a linguistic difference in the vowels, most likely length.

The 'every-day' phonetic syllabification I have proposed as the phonetic basis for the evening out of the length of stressed syllables was then different from another phonetic syllabification used in poetry. The every-day maximalistic syllabification can be assumed to have gradually gained ground, causing the old short vowels to lengthen. We could then perhaps just as easily ascribe the cause of the metrical change in the 16th century to an increased prominence of the new stressed syllable, as to increased length of the old short vowels. Similarly, the relative conservatism in the last feet of untruncated lines can be looked on as a consequence of the resistance of the old 'poetic syllabification' which did not produce stretchable syllables from the underlying forms of words like tala.

2.4 Conclusion

In the foregoing I have been trying to formulate a description of the quantity situation before and after the quantity shift in terms of syllabification. If these hypotheses bear some relation to the truth it becomes natural to look upon the quantity shift as being basically the rise of the final-maximalistic stressed

syllable as a central unit in the phonology of Icelandic.

I am suggesting, in other words, that the basic difference between the prosodic system of Old Icelandic and Modern Icelandic is that in the latter the stressed syllable plays a central role in the phonology.

The final outcome was, then, that the length of vowels became (via the stressed syllable) defined by the stress pattern. Before the shift, length was a paradigmatic feature, and long or short vowels could more or less freely combine with a long or a short following consonantism without any limitations imposed by stress or syllabification. I have suggested that phonetic stress was the cause of the changes, both the lengthening and the shortening, although in a slightly different way for the two changes.

As we have already said, we don't have to assume that the change took place in one big leap from the Old Icelandic system to the Modern Icelandic one. There probably was a long period of instability, but when the phonetic alternations had become regular enough, a phonological reorganization could take place as an innovation made by new speakers confronted with the output of their seniors showing, on the surface, data which could be accounted for by rules of the type described in Chapter II. For the words with *p, t, k, s + y, j, r* following the vowel, a final-maximalistic syllabification triggered by the stressedness of the syllable ran up against a constraint which prohibited syllable-final

sequences where the consonants v, i, r followed p, t, k, s, so the syllable was only extended far enough back to include the 'strong' consonants p, t, k, s. This meant that for the forms concerned the domain of the length rule only came to be a vowel followed by one consonant, so the vowel became long.

As to when all this took place, we have suggested that the phonetic change had begun in the 16th century, but the fact that poetry from the latter half of the 17th century and even from the 18th century shows discrimination between old long and short 'syllables' seems to indicate that the new system was not stabilized until much later, perhaps as late as the beginning of the 18th century. It is furthermore quite likely that the change progressed with different speed in different geographical areas.

It seems natural for a change like this one to progress gradually, since it did not lead to any clashes in the system. It seems that as early as the 13th century the long and short vowel systems started developing separately, so that qualitative differences gradually took over the distinctive function formerly carried by the length. Once these qualitative differences had developed, the road was clear for a reorganization of the prosody. But there was no hurry either; only perhaps a relatively unstable tendency acted to bring about the change.

Yet, when the change was completed and order restored to the former 'chaos' concerning the length of stressed segments and syllables, it can be called a major change in the history of Icelandic.

3. Explaining the quantity shift.

In this section we will turn our attention more closely to the problem of what caused the changes we have been looking at (Question 3. on p.196). A question like: Why did change X take place? may seem stupid in the context of this thesis, to judge by often quoted remarks by prominent linguists: "the causes of sound change are unknown" (Bloomfield 1933:385); "The explanation of the cause of language change is far beyond the reach of any theory ever advanced" (Harris 1969: 550). Such pessimistic remarks may lead one to wonder whether the question we have just posed is worth asking at all, since it would seem that we are far from being able to answer it. But a question which is never asked seems to be rather unlikely to be answered, and if we hope sometime to be able to say something about the causes of particular changes and make statements about the causes of language change in general, we will obviously have to put some effort into trying to find at least tentative answers to questions like the one we have just put forth.

It is useful to start out by considering what sort of a sensible answer we can expect to get to our question.

We will, for example, have to form as clear an idea as we can about what we can possibly mean by the notion 'cause of a linguistic change'. Closely related to this, we will have to consider what is meant by such a phrase as a 'valid explanation in historical linguistics'. We will have to have some ideas about what conditions a piece of historical linguistics must fulfil in order that we can call it a valid explanation, since, presumably, a part of an explanation of some specific historical change will be a statement about its causes.

3.1 Explanation

I would like to start by making some comments on the term 'explanation' in linguistics in general and later turn to its possible uses in historical linguistics. In recent years questions of methodology and the nature of explanations in synchronic linguistics have received a great deal of attention. (See e.g. Cohen (ed.) 1974, Derwing 1973, Botha 1971, Lass 1976a (especially pp. 213-220), and Langendoen 1976). The reason why these matters have been the cause of so much paperwork is that, probably more than in most other subjects, it can be a matter of dispute, not only how to explain linguistic phenomena synchronically, that is, what sort of theoretical machinery is needed, but it is also a matter of dispute what there is to explain. Obviously, there is no room here for a thorough and sensible discussion of these matters, since the confusion seems to

be almost hopeless. But I will try to clarify some points which may be relevant to what I have to say about Icelandic.

One question which may be relevant is what we might call the question of levels of explanation. Chomsky (1965) distinguishes between what he calls 'descriptive' and 'explanatory adequacy'. This distinction has to do with the dichotomy between universal and particular grammar (on the psychological side: 'linguistic competence' and 'faculté de langage'). It seems to me that Chomsky's point is that a theory is 'explanatorily adequate' in so far as it 'explains' facts about particular grammars in terms of a universal grammar. This implies that linguists are concerned with at least two types of activities, (a) describing particular languages by writing grammars for them, and (b) describing what are the common features of all grammars by writing a Universal Grammar and relating it to particular grammars. If we look at this in terms of explanation, we could say that Chomsky's point is that facts about particular languages are explained by particular grammars, but facts about particular grammars are in turn, in some sense, explained by Universal Grammar. We seem, then, to be able to talk about explanations at at least two levels, that is one can explain facts about languages by writing particular grammars and one can explain facts about grammars by writing Universal Grammar. (Obviously there is an interesting, and maybe quite uncomfortable, question

lurking behind this one: What explains Universal Grammar?)

Another question is of course the question of what is an explanation in linguistics. I will deal with this by trying to form some idea as to what, in general, could be called an explanation of some phenomenon and consider how things we might want to call explanations in the context of linguistics compare to such a general notion. I will not try to give a review of the literature on this very central problem in the philosophy of science, but rather make an attempt to express my intuitions about matters related to it which to some extent are influenced (in a positive or negative way) by my reading of e.g. Bach (1974), Lass (1976a:213-220) and parts of Popper (1968). I would in no way want to maintain that what I have to say solves any problems, it is mainly meant to put what I have to say about Icelandic in some (to me) sensible ontological context.

In a sense, one can say that an explanation is simply a higher order description of any event or state of affairs. This seems to make some sense if we look at a simple example. We may observe a simple fact, say, that John has a pain in his toe. We have described a state of affairs. We can find an 'explanation' for that state of affairs by observing the fact that there is a splinter stuck in John's toe. Evidently, this 'explanation' is simply a description of another state of affairs, that state of affairs being that a splinter in

John's toe irritates his nerves so that he feels pain. If we can believe that the presence of the splinter is related to the pain, we may accept this as a valid explanation. We might like to stop there and say that we have explained the fact that John has a pain in his toe, but we may choose to be more curious and want to explain this more deeply. Here, we can go in different directions. We might want a more detailed account of the relation between the fact that there is a splinter in John's toe and the fact that he feels pain. In that case we might enter on the activity practised by neurologists, who try to describe in more detail how irritation of nerves causes sensation of pain. We might not be interested in that problem, however, but rather we might want to explain the state of affairs that John has a splinter in his toe. We could do that for example by observing that he just walked barefoot on a wooden floor. We can describe that fact and in that way 'explain' the state of affairs that John has a splinter in his toe. That might arouse our interest in knowing more about how the splinter got into John's toe by describing in more detail what sort of movements John made when he walked on the floor, and what state the floor was in. In so doing we might observe some facts that still arouse our curiosity and carry on finding new explanations, which turn out to be simply descriptions of some new states of affairs. Each state of affairs can be said to be explained by a more penetrating description in which new facts are brought out. In this

definition, then, a description of facts on a deeper level of observation, call it level n , is an explanation of facts observed at a more shallow level, level $n - 1$. Each explanation gives rise to new puzzles, and this can go on indefinitely, unless perhaps ultimately we could reach a state where we have explained all phenomena, in which case we might have reached God, which (who) is perhaps a state of affairs which (who) cannot be explained by another state of affairs. In the case where we are not looking for God, we might only want to go a certain distance on this infinite (or ending in God) ladder of explanations; the limit at which we stop is determined by things like our intellectual capacities, the amount of time we have on our hands and the degree to which our curiosity is aroused. We may stop at the observation that John has a splinter in his toe and just take that for granted, or we might be more curious, which could, given enough time, energy and ingenuity, lead us to the forefronts of several sciences (neurophysiology, physics etc.).

To go back to Chomsky's descriptive and explanatory adequacies, we could, as I said before, say that a grammar (for Chomsky meaning also 'competence') for a particular language, being a description of the system (or lack of system) behind the sentences of that language 'explains' the corpus of the sentences of that language. It is, in Chomsky's terms 'a theory of the language'. On a deeper level, Universal Grammar (being

'systematically ambiguous', meaning also faculté de langage) can be said to explain particular grammars, that is, it can tell us, in part at least, why the grammars of particular languages are the way they are. In describing Universal Grammar we find a basis for (perhaps only partial) explanations of particular grammars. If we are curious enough, we could of course ask ourselves why Universal Grammar is as it is. This is an interesting question, but we will not discuss it here for obvious reasons.

Before trying to clarify in more detail what we would like to call an explanation in linguistics, I have to touch on a third problem which comes up in this context, that is, what there is to be explained. This is really the question of what the domain of linguistics is. Some would want to say that linguists should explain, as Langendoen (1976:690) puts it, 'all systematic linguistic phenomena, including those dealing with the use of language'; others would want to say that the scope of linguistics is more narrow, that only some regularities observed in linguistic behaviour should be dealt with (the 'narrow scope' view, cf. Langendoen loc. cit.). It is for example possible to study what people say without studying why or in what context they say it. Another question pertaining to the scope of linguistics is the question of psychological reality, that is, whether linguists describe or should describe entities (unconsciously) present in people's minds or whether they should,

or do, describe some systems existing at more abstract levels outside individual speakers. Proponents of the latter view maintain that the accounts they give of linguistic regularities and irregularities are not necessarily (nor have to be) psychologically real in the sense that they are isomorphous with psychologically present linguistic systems, which people use (among other things) when they speak. Obviously, there is no room here for a detailed discussion of these matters. In fact we can form some sort of an idea of what we could call an explanation in linguistics without having solved the question of the domain of linguistics.

We don't have to take a stand as to mentalism or non-mentalism or to the question of how many aspects of people's linguistic behaviour linguists should explain in order to form an idea of the concept 'explanation in linguistics', although, of course, particular instances of potential explanations will not be sorted out as valid or invalid unless we have some idea as to what we want to explain. It is sufficient for our present purposes to stipulate that synchronic linguistics is dealing with some sort of reality. We can, for the moment, stay neutral as to what this reality is, whether it is the competence of an ideal speaker-hearer or a socially real set of rules and paradigms, or an even more abstract system, having an ontological status similar to logic or mathematics. But it seems essential that some reality is assumed for these things, since if they are not real,

it is in principle impossible to test the validity of proposed explanations, if the concept explanation is thought of in the way I have just suggested.

If an explanation of something is a description on a higher level, there must be something to describe. It is, then, incompatible with my concept of explanation to take what may seem to be the relatively modest stand (and therefore more sensible at this stage of our ignorance) to propose grammars for languages which account for, or generate, all the forms of the language (e.g. enumerating recursively all the sentences and giving them structural descriptions, making statements about the syntactic, semantic and phonological structure of the sentences) without claiming that the derivations and structural descriptions arrived at by the grammars represent any reality, that is, assume that the grammar is scientific fiction. This is incompatible with my notion of explanation, since in this way grammars will be in principle untestable and we will never have a way of assessing the validity of their explanations. A realist grammar, on the other hand, purports to describe some reality, and in principle all that has to be done in checking its proposed explanations is to compare them with that reality to see whether they describe the actual state of affairs or not.

A realist grammar, then, is empirical in a very general sense, there being assumed some reality against which it is to be tested. (It is, of course, a different

matter whether it is in fact possible to test the proposed grammars against the reality they are claimed to represent. I will come to this shortly.)

A non-realist grammar has a very different status. The only condition it has to fulfil is that it fit the corpus of linguistic surface data available to direct observation, that is, that its predictions are not contradicted by actual sentences produced by the speakers of the language under investigation. It has proved to be the case that within this framework there are more often than not many ways of accounting for some piece of linguistic data. Different grammars can generate the same strings. It is thus possible to get more than one explanation of a particular linguistic phenomenon which fits the data, and if we ask which explanation is the correct one, or only the best one, granting that perhaps none is perfect, we seem to be in trouble, because it seems to me that we can't give the word best any reference if we are not making claims about a reality. One might suggest that the 'explanations' should be tested against an evaluation measure of some sort which would be built into the theory in order to tell which explanations are the best ones. But evidently the same problem comes up concerning the evaluation measure, it becomes vacuous if no reality is assumed. If there is no external reality in terms of which the evaluation measure can be defined, it can only be defined in terms of the system within which it is supposed to operate, for example by working out a simplicity metric. One might for example suggest that

the explanation which uses fewest symbols (presumably defined in some clear way) is the simplest and therefore the best one. But different theories will have different primitive concepts, and there is no guarantee that an explanation which is simple in one formalized theory will be simple in another, if translatable into that theory at all. And if different theories choose different explanations, we are faced with the task of choosing between different theories if we still want to have one rather than many valid explanations. But choosing between theories poses exactly the same problems as choosing between different explanations within a theory. We will have to ask on what grounds we can evaluate different theories. If this is possible at all, it would have to be done within a metatheory which incorporates an evaluation measure which distinguishes between different theories. But it is difficult to see how such a metatheory can be chosen other than by a still higher evaluation measure, incorporated into a meta-meta theory, because how else do we know if we have picked the right way of choosing between different theories? Evidently, this evaluation hierarchy is infinite, if we don't have some real referent to test our theories and evaluation metrics against. What this shows is that non-realist synchronic linguistics, being untestable against reality, can't decide what is a good or a bad explanation of something. ¹⁰⁾ So, if we want to be able to evaluate our explanations, it seems that we should stick to some sort of realism.

Before moving on to historical linguistics, let me make a note on the testing of hypotheses within our realist framework. In this framework the term 'valid explanation of a phenomenon X' will mean something like 'a true description of the state of affairs which can be said to have brought about this phenomenon' or 'a true description of the context in which the phenomenon exists'. But then of course the question arises how we know that we have given a correct description of the phenomena. More often than not, the 'deeper states of affairs', which we describe and connect with the more observable phenomena which we want to explain, are not directly observable. But how can we, then, know whether we have described them correctly, that is, whether our explanations are valid or not? According to Popper we will perhaps never know whether our descriptions are true. The best we can do is to limit the number of descriptions that may be true: "Only the falsity of the theory can be inferred from empirical evidence ..." (Popper 1968:55. Emphasis his.). This can be done, according to Popper, by testing the predictions made by the proposed theories. If we have proposed some unobservable state of affairs to prevail, it may have some more or less logical consequences in that it predicts some other phenomenon which may be observable. If these predictions are contradicted by observable facts, our theory is wrong. Popper suggests that the difference between scientific explanations and non-scientific

ones is that the scientific ones must be refutable. Irrefutable explanations and theories are unscientific, according to Popper. For an explanation to be refutable by observable phenomena, it will have to be formulated in such a way that its empirical consequences are clear in that predictions about some phenomena which should be observable can be deduced from it in some logical way. This is closely related to the concept of law; the explanations will have to be predictive laws. (The laws being, in our terms, quantified descriptions of the relations between phenomena or states of affairs. - The iff relation is a statement that every time one thing occurs, another must also ^{have} occurred.)

If we go back to the example of the splinter in John's toe and his sensation of pain, the explanation we proposed for the fact that John had a pain in his toe was the fact that he had a splinter in his toe. This is not a law, it is only a hypothesis about a particular state of affairs or event, it does not have the all-quantifier in it, it does not make predictions. In order to make this explanation refutable à la Popper it would have to be turned into something like this:

"Whenever a person has a splinter in his toe, he or she feels pain." In order to test the validity of this law we would go out and look for counterexamples to it, and if we find a person who has a splinter in his toe, but still does not feel any pain, our law is refuted. In Popper's framework, one instance of a person with a

splinter in his toe but without pain invalidates the law. But the important question is: Does this mean that our explanation of the pain in John's case is invalid? I should think not. In explaining the fact that John has a pain in his toe by the fact that he has a splinter in his toe, we assumed that there was a connection between the two facts, so that John would not have had the pain if the splinter had not been there. This may or may not be true. But the truth of that statement about John's case does not depend on whether we find some other person who does not have a pain although he has a splinter in his toe. So, if we take a valid explanation to be a true statement about a state of affairs, the falsity of that statement, which is made about one state of affairs, cannot be proved by showing that it does not hold for another state of affairs, perhaps similar, but slightly different. The point is that our explanation is not a predictive law. We may be right in assuming that there was a connection between John's pain and John's splinter without making a predictive statement that every time a person has a splinter in his toe, he will feel pain.

As mentioned above, Popper suggests that for something to qualify as a science, it must form refutable theories, that is, theories that make general predictions about things other than those observed. It seems to me that it is not right to put such a strict constraint on the definition of science. I think that there is more to science than mere theory-making and theory-refuting.

Description of what are assumed to be the facts seems to me to be a quite legitimate scientific activity, if not more important than theory-making. It may well turn out to be the case that our descriptions of simple facts, basically, have a common core with predictive general theories. — The particular descriptions are in essence hypotheses about particular states of affairs. (This is definitely Popper's view, I think, cf. Popper 1968:27-29). But there is a very important difference between a hypothesis about some particular state of affairs and a predictive law. Predictive laws claim that the descriptions they incorporate are valid for more cases than the ones that they are originally applied to. But it is impractical to form laws unless we can be pretty sure that we have taken all relevant factors into consideration. There may be a connection between the splinter and the pain, but there may also be other conditions necessary for a person to feel the pain. To admit that there may be other things at work is not to say that it is completely useless from the scientific point of view to make the assertion that there is a connection between the splinter and the pain. I would say that it is perfectly sensible to say that the fact that John has a pain is 'explained' in a technical sense by the fact that he has a splinter in his toe and that we can call this a 'scientific explanation' (although it is admittedly not a very 'deep' one). But explanation in this sense is evidently not a law, and that means that the validity

of our explanations cannot be tested by Popper's method. There may thus seem to be a conflict between my concept of a scientific explanation and Popper's criterion for scientificity. My explanations are not refutable, simply because they are not predictive laws, but I would maintain that they are empirical claims and in theory testable against the real state of affairs. But in order to be able to test the explanations, one will ultimately have to 'see' the real states of affairs, to compare the theories to that reality, and in many cases our perceptual apparatus, both natural and man-made, is not up to the task of making these direct observations of reality. In the absence of direct observations we can only make more or less plausible guesses, but it is only when we move from the stage of claims about particular, more or less isolated pieces of reality which can be said to be true or false hypotheses about particular phenomena, to the stage of forming predictive laws which can be falsified by counterexamples, that Popper's method can help us. It seems to be an unreasonable puritanism to maintain that only the quantified laws are scientific statements.

We will now turn to historical linguistics and consider what could be called a 'valid explanation' in that principle. This question is obviously related to the main theme of this section, namely the causes of the changes we are studying. If we can find something which we can call the cause of a linguistic change, it seems

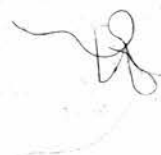
reasonable to say that that forms at least a part of an explanation of that change. Before looking at the question of cause, we will, then, discuss briefly the general question of what can be called a valid explanation in historical linguistics. We have already seen that the question^{of} what counts as a valid explanation is closely connected with the testability of the truth of explanations. Hypotheses in a mentalistic synchronic grammar or a grammar purporting to explain some social reality, we said, should in theory be testable against the presupposed reality. We saw also that a 'non-realist' synchronic theory did not seem to be testable against any reality except a corpus of sentences produced by the speakers, which by itself seems to be of rather limited value, if no recourse is allowed to any hypotheses or assumptions about the reality of linguistic rules. In order to put historical linguistics in the same perspective, we will consider against what sort of reality hypotheses about language change should be tested or in general whether they can be tested against any reality, that is, whether historical linguistics is an empirical science.

In discussing the status of synchronic hypotheses, we suggested that unless some reality was claimed, synchronic linguistics would stand in the void, since its hypotheses would be untestable. The reality that must be assumed can either be a mental reality in individual speakers, or a social reality, that is, synchronic linguists can claim to describe the competences of speakers or

the linguistic systems of communities. (There is also the possibility of assuming a more abstract ontological status for linguistic systems, but we will leave that out of the discussion for the moment.) It may be equally difficult to test hypotheses about social reality and hypotheses about mental reality, but it is theoretically possible in both cases, some sort of a reality being claimed to exist and the aim being to describe it.

As to what sort of a reality historical linguistics should aim at describing, it is evident that it could not be simply defined as mental reality of speakers. It is for example difficult to see how such a thing as Grimm's Law can ever have in any sense been mentally real to any speaker. If historical linguists are committed to mental realism, they will obviously have to look, not at the competences of speakers in isolation, but rather compare the competences of speakers from different times and describe the differences between them. But it has been argued, most notably by Weinreich, Labov and Herzog (1968) that this is not enough. They argue (p. 156) that one should "abandon the individual homogeneous idiolect as a model of language" in order to be able to "suggest a more intelligible mechanism of transfer [of linguistic innovations]". What seems to me to be the main thesis of their paper is that if linguists are to understand language change, they cannot limit their scope to the competences of individual speakers, but have to be able to look at the linguistic systems of communities and to

study variation between speakers, for example according to age or social stratification. Not only this, the theory will have to allow for a conception of competence, not as homogeneous, but allowing for variability, for example accounting for switches speakers often make between styles, according to context, etc. So, part of what historical linguists have to do is to describe language systems of social groups and individual competences, seen as fluctuating systems, and they have to compare different language systems, both social and individual ones, from different times and places. But even though a historical linguist could make true statements about different language systems and describe the differences between them, that would not be enough. For example, statements like: "In Old Icelandic, vowel length was phonemic, but became context determined in the 16th or 17th century", seem to lack something. What statements of this sort do, is merely to describe correspondences between two stages in the history of the language in question: Where we have X at stage A we have Y at stage B. Even a sophisticated model like that proposed by Weinreich, Labov and Herzog is not enough, if it merely compares one fluctuating system with another. Historical linguists have claimed that for something to be an explanation in historical linguistics, it is not enough to describe the situation before and after some change, but that it has to say something more about the relation between the two stages (Cf. e.g. Jeffers 1974:236 and



H. Andersen 1973). For example, we want to know why the change took place, and also how it took place. This is where the concept 'cause of a linguistic change' comes into the picture.

The most obvious way to look for what we might term as causes of a change of some linguistic form is to consider closely the context in which the form occurs before or at the time the change takes place. This context could be not only the purely linguistic or systematic context of the form, but also, as stressed by Labov and his collaborators, the wider social and cultural context. Still another part of the context of a linguistic form is, of course, the phonetic medium, describable in terms of articulatory processes and acoustic laws. This last type of context is of particular importance when phonological changes are studied. If all these surroundings of a linguistic form which has undergone a change are studied closely, we may be able to find some factors which we could suggest, in some sense, caused the change, or explain why it took place.

In our comments on the notion 'explanation in historical linguistics' so far, we have assumed that there is some reality which the historical linguists should be trying to describe, that is, we have assumed that the discipline is in some sense an empirical one. But one may well wonder whether this is necessarily so, that is, just as, theoretically, it is possible to write synchronic grammars without making any claims about psychological

reality or social reality, it would seem possible to look upon historical linguistics in the same way. One might then say that when there is no external or empirical evidence about what happened, one will choose a descriptive model without claiming that it represents the facts. The model might then be chosen on grounds of simplicity defined in terms of the formalism used. But obviously, this would lead to the same problems as we have seen that would come up in synchronic linguistics, as far as the limitation of plausible explanations is concerned. Another fact which makes it less plausible to think of historical linguistics as a metaphysical discipline is that it seems, perhaps somewhat paradoxically, that facts concerning historical changes are often more observable than facts about unconscious mental processes and structures. Indeed, it seems to me that it is always assumed by historical linguists that they are trying to find out what really happened and the things they are trying to describe are, to them at least, in some sense real. I think a historical linguist would get a funny look from his colleagues (not to mention other more normal people) if he were to say: "I am not interested to know what really happened in the history of language X, I am only wondering what is the most beautiful (= simple) picture of what might have happened". What our imaginary historical linguist just said seems to me to represent fairly what would be the point of view of historical linguistics which does not claim to

be trying to describe some reality, in other words, a non-realist historical linguist. (I am not, needless to say, claiming that considerations of simplicity are illegitimate or useless in historical linguistics. Indeed, it is considerations of simplicity that prevent us from postulating events that there is no evidence for. We would not, for example want to explain a correspondence between [a] at stage A and [æ] at stage B as representing two changes [a] → [e] and [e] → [æ], if there is no external evidence to the effect that this happened. Also, simplicity is at the center of old established conventions in historical linguistics like the comparative method. The use of the simplicity arguments in these instances is based on the policy that until ^{wrong} proven it is best to assume that things happened in a simple way.)

If we now try to summarize what we have been saying about what would qualify as an explanation in historical linguistics, we seem to come up with a rather complicated concept. An explanation in historical linguistics will have to make claims about at least two different synchronically defined stages in the history of some language, that is, the linguistic systems of communities and/or competences of individual speakers at these stages; it will also have to make claims about correspondences between the systems at the two stages; it will also have to make claims about the relation of particular forms at the first stage which show changed

reflexes at the second stage to factors in the context of those forms in a way that we can say that the context in some sense explains why the change took place. Further, the explanation may have to make more specific claims about how the change took place, in other words how the 'cause' is related to the change and the result.

We will assume that what was said in the preceding paragraph constitutes at least a part of a definition of the concept 'explanation in historical linguistics'. This seems complicated enough, but we still have to state what we would like to call a 'valid explanation in historical linguistics'. As in synchronic linguistics, we seem to have to do this in terms of reality and truth. In order to be able to test hypotheses in historical linguistics, we will have to assume that there is some reality we are trying to describe. The linguistic systems of communities or individuals must somehow be real, and the changes must also be real and have taken place for specific, perhaps complex, reasons, and in specific, perhaps complex, ways, and two contradictory explanations cannot be true. We will thus, in principle have to test different explanations against the reality we are trying to describe. As in synchronic linguistics the testing of hypotheses against the reality can be very difficult or impossible in practice, but in theory it is possible, if the reality is assumed to exist. If we don't do this, there seems to be no theoretically possible way of testing the validity of proposed explanations.

3.2 The term 'cause' in historical linguistics

We will now try to get closer to what was supposed to be the main theme of this section, namely the cause, or causes **of** the quantity shift. As a preliminary to that, we will still have to make some general remarks, this time about the concept 'cause of a linguistic change'.

Closely related to the term cause, is the term condition. The terms 'cause of a change' and 'condition for a change' are often used to refer to the same thing. It seems to make some sense to say something like this: "Under condition X, change Y will take place", or "Under condition X, change Y can take place". In statements like these it is assumed that there is some connection between the condition X and the consequence, which is the change Y. We note that in the two conditional statements above, different claims are made. In the first one, it is claimed that, given the condition X, the change Y will take place, but in the second, it is only claimed that, given the condition, the change can or may take place. In the latter it is only stated that the condition X is a necessary condition for the change, but it is not assumed that the change will take place. In the former statement, on the other hand, it is maintained that the condition X is both necessary and sufficient for the change to take place: Given the condition, the change both can and will take place.

If we now put the term 'cause', to be used in historical linguistics, into the perspective of necessary and sufficient conditions, it seems to me that it can only be used in a sense similar to that of 'necessary condition', and that we will never find sufficient conditions for linguistic changes. I am not saying, of course, that linguistic changes don't have sufficient conditions. The fact that changes occur must mean that, in some sense, there are sufficient conditions for their taking place ('sufficient condition' meaning in this context the surroundings that make it necessary for the change to take place). I am only maintaining that it would be too much to expect linguists to find all these sufficient conditions for all linguistic changes.

If it were true that for any change X it is possible to find sufficient conditions for it (i.e. there is no change Y such that it is theoretically impossible to find sufficient conditions for it), that would mean that it is theoretically possible to find sufficient conditions for all changes a given language is going to undergo. If there are no limits that discriminate against any language as far as this is concerned, it would be possible to find sufficient conditions for every single change that every single language is going to undergo. Thus, if it is in principle possible to find sufficient conditions for linguistic changes, it means that it is theoretically possible to predict every single change that every single language is going to undergo from here to eternity.

We are here faced with a problem common to all evolutionary sciences (evolutionary biology, genetics, history etc.), the problem of whether evolution can be predicted (cf. Scriven 1959). The question can be dealt with from two points of view. One can wonder whether it is logically possible to predict evolution. This is a philosophical problem, the answer to which probably depends on whether something like complete randomness exists. We will not try to solve this here. The other point of view is the more practical point of view of individual disciplines. In historical linguistics, as well as evolutionary biology, it is not only linguistic and biological factors which determine how linguistic or biological entities evolve. There are always present other 'external' factors which mess things up: "... the irregularity-producing factors lie outside their [i.e. the disciplines'] range of observation and are not predictable by reference to any factors within this range" (Scriven 1959:478). For language, these external factors are for example physical surroundings (it is possible for a language to die out because all its speakers die in an earthquake or some other natural catastrophe), but probably the most important 'irregularity-producing factor', as far as the evolution of language is concerned, is the human will (sometimes harnessed into currents of social laws and etiquettes), which is in principle unpredictable from the point of view of linguistics. If historical linguistics

is to be a branch of linguistics, working with more or less the same technical apparatus as other branches of linguistics, it is impossible to demand from it that it find sufficient conditions for linguistic changes and by that make them predictable. If linguistic changes are to be predicted, it requires knowledge of all sorts of things which have nothing to do with linguistics; in fact, historical linguistics would have to become a sort of theory of everything.

It seems, then, that if the term 'cause' is to have any meaning within historical linguistics, it certainly can't mean 'sufficient condition'. It will have to mean something similar to the 'logical' concept 'necessary condition'. In dealing with historical change, one can often find certain features which seem to be related to those features that change. For example, it has been noted that, in Germanic, stem-vowels were fronted in forms where an i or a j followed the stem. The presence of a following i or j was a condition for the fronting of the vowel. Evidently, we are not talking about a sufficient condition, in that whenever a stressed vowel is followed by an i or a j (with some intervening consonants) it gets fronted. There are exceptions to this, not only in that in many languages i's and j's can follow stressed vowels without fronting them, but also in that there are quite a few exceptions to this in Germanic: Old Icelandic staðr 'place' (< *staðiz) does not have a front e in its stem even though it was followed by an i

at the time when the fronting of the stem vowel of gestr (< *gastiz) took place. Many of these exceptions can be explained by more or less general features which interfere with the otherwise valid law, and the fronting did not have to take place. But the important thing is that we can say that if the i's or j's had not been there, the fronting would not have taken place. The presence of i's and j's was a necessary condition for fronting; there was a relation between the fronting and the presence of the i's and j's. It seems perfectly reasonable to use the term 'causal relation' to denote this. We can thus say that i's and j's caused fronting of preceding stem-vowels.

So far, so good, but we have not done enough. If we simply state our definition of cause in terms ^{of} the concept 'necessary condition', we seem to have a pretty wide definition. We may, for example, observe that a language won't change unless it is spoken (or used). Thus we can say that if language X had not been spoken, change Y would not have taken place. It is a necessary condition for change Y that the language X was spoken, and thus, according to our definition, the change Y was caused by the fact that X was in use. Of course this can be said to be trivially true, but no linguist would consider it a great achievement to be able to state this. Our definition of cause may thus lead to such, in a sense, absurd uses of the concept. But that may not be a bad thing. In reality, this is exactly the same

thing as we saw that cropped up in our discussion of explanation. Explanations can be valid, but trivial. The explanation using the fact that language X was spoken to explain a change in that language is one of that type; it is perfectly legitimate, but it is very uninteresting. We are not surprised or enlightened by this observation. It is only when we are surprised, when we experience some sort of revelation when faced with an explanation or a discovery of a cause (or causes) that we feel that our effort has been worthwhile. We can thus say that causes are more or less obvious and explanations more or less interesting. Since our investigation is generated by curiosity, we will not want to discover what we already know, but rather look for new answers. Thus the number of things that will be put forth in research will be greatly limited by this attitude of the researcher and his colleagues. The more interesting the discoveries the better.

This, however, is not enough to limit the use of the term cause in a way that seems desirable. It is not only that we want to be surprised by good explanations, we want to be sure that they are 'relevant' (cf. Hempel 1966:48). We would like to eliminate explanations like: "The quantity shift was caused by the execution of Bishop Jón Arason" ¹¹). The question is how we can do this. The claim made by this explanation, referring it to our definition of cause, is that if Jón Arason had not been executed, the quantity shift would not have taken place.

But, of course this can't be tested; we will never know what would or would not have happened if Jón Arason had not been executed. So, on the grounds of our definition we can't get rid of the hypothesis claiming that Jón Arason's death caused the quantity shift. This is, of course, intolerable. The common-sensical answer would be that the death of Jón Arason had nothing to do with the quantity shift. This is really an empirical claim. It is claimed that in reality, the nature of Jón Arason's death and the nature of the quantity shift were such that they happened independently of each other. And when we connect the fronting of stem vowels in Germanic with the presence of i's and j's we claim that the nature of the fronting and the nature of the i's and j's and their relation to the stem vowels were such that there was a connection. This is either true or not true. We find it likely to be true on various grounds. Our belief that this is so makes it sensible to claim that the i's and j's caused the fronting of stressed vowels.

It must be admitted that the concert of cause that is intended to emerge from this somewhat lengthy digression is far from being precise. Indeed, it has been maintained by philosophers like Hume that the causal relation exists only in the minds of people and is an, often misguided, interpretation of the relation or non-relation between two events (cf. von Mises 1951:151-162). But I have a feeling that there is more to the causal relation than that, that in reality there are some events

that are connected to some other events in a way they are not connected to still other ones, and to express this, human language, scientific as well as every day, needs some concept.

In the belief that this is so and on the basis of the 'clarification' of the concept cause, suggested here, I will carry on happily in spite of the pessimistic comments quoted at the outset of this section, implying that there is little hope of finding the causes of linguistic changes. If the meaning of the term 'cause' as a linguistic terminus technicus is restricted along the lines described above, there seems to be nothing wrong with applying it in historical descriptions.

3.3 The causes of lengthening and shortening.

Let us now at last turn to our particular problem and start wondering about the cause, or causes of the quantity shift. We have already suggested that it is not necessary to assume that the shortening of old long vowels in front of two or more consonants took place at the same time as the lengthening of short vowels in front of no more than one consonant. Similarly, it is not necessary to assume a priori that the two changes had the same causes.

As far as the shortening of old long vowels is concerned, we have suggested that at an early stage there was a phonetic tendency to have stressed underlyingly

long vowels shorter than directly predicted by their underlying features when they were followed by two or more consonants. We tried to make this plausible by assuming that there existed an articulatory unit, we can call it the phonetic syllable, which was delimited by stress. This unit, which we can assume was primarily a phonetic thing, occurring in actual speech performance, was not necessarily definable in terms of the linguistic system, nor was it necessarily a structural unit (for example in that it figured in any systematic rules of the phonology). (Presumably, unstressed phonetic syllables were different performance units, and probably different phonetic laws applied there. We may leave them out of the discussion here, since we are only concerned with changes that took place in stressed syllables.)

Having assumed a phonetic unit which we call the stressed syllable, we will assume further that there was a tendency for the units to be of approximately the same duration. Given these assumptions, we can easily imagine that when a phonological form like fátt, 'few' (neuter), one syllable with a phonologically long vowel followed by a geminate, which was presumably relatively long in its phonetic duration, was pronounced as a phonetic, stressed syllable, its vowel tended to be shorter than the phonologically same vowel of the form fát 'confusion' which was followed by only one consonant. We can thus imagine that the shortening of long vowels was originally initiated by a tendency of the long vowels to adapt to their

phonetic surroundings. These phonetic surroundings were the stress and the following consonantism. The phonetic stress delimited the immediate context of the vowel, and when a part of the context was a long following consonantism (a geminate or a consonant cluster) the vowel was relatively short. At first, this shortening of the vowel was probably irregular and varied according to speakers, some speakers shortening the vowel more often and more regularly than others. This can well have varied according to dialects or class, and it should be easy to fit this sort of thing into a model similar to the one used by Labov (1965) and Weinreich, Labov and Herzog (1968) for variation in American dialects.

At a later stage we can assume that the shortening of long vowels became more regular, until some language learners picked it up as a full-fledged phonological rule of the language and made the generalization that all stressed vowels are short when occurring in front of two or more consonants. It is obviously a difficult problem to decide exactly at what stage a new rule like this becomes, so to speak, an integral part of the phonology, so that all or the majority of the speakers have it as a regular phonological rule, as opposed to a more or less irregular phonetic alternation, and indeed it may be that it is theoretically unsound to make such a distinction, since it would seem that borderline cases will always come up, where it is impossible to decide whether to call something a low-level phonetic alternation or a regular

phonological process. One might perhaps say that the problem is a pseudo-problem, only created by an unjustified dichotomy between phonetics and phonology. But it seems necessary, at least for practical purposes, to draw some distinction between completely phonetic alternations, conditioned, e.g. by the capacities of the speech organs or some physical laws of acoustics, and systematic phonological alternations, defined by the rules of the grammar, phonological rules which are language specific and learned by speakers when they are learning the language, but not explainable in general phonetic terms.

In this context one may ask, from the point of view of historical linguistics, what in the process we have been describing above, we should call the linguistic change. Is it the appearance of the phonetic tendency to shorten long vowels in certain phonetic environments, or is it the actual incorporation of the phonological rule into the language system, that is to say, the change in the 'competences' of speakers of the language? We are evidently faced with a situation which is reminiscent of historical structural linguistics, where there is a distinction between an allophonic change and the restructuring of the phonological system, the phonemicization of a previously allophonic difference. It was common among structuralists to claim that the only thing that mattered was the structural change: "Phonetic change acquires significance only if it results in a change of

the phonemic pattern." (Bloomfield 1933:369). This may be true to a certain extent, but as soon as it comes to asking questions of how and why (instead of only what) the phonetic drift becomes extremely important. If we only bother about structural changes, we may in fact end up comparing different synchronic stages without being able to explain their relation. A historical linguist should be just as interested in the phonetic aspects as the structural aspects of changes, and therefore there does not seem to be any justification for giving one aspect more priority than the other by singling it out as the change proper.

We have now set forth a hypothesis concerning the cause and nature of the shortening of long vowels before two or more consonants. We could say that we have proposed at least a part of an explanation of that part of the quantity shift. It is now reasonable to ask, on what grounds we may think that it is valid, that is, do we have some evidence to support it? Part of the justification of our hypothesis is implicit in our description of it, which we have given above. We tried to give the explanation some plausibility by making it believable from the phonetic point of view. But we may ask: What else do we have? The answer to that question is: Very little. As we have said before, the testing of hypotheses like the one we have proposed is an empirical matter. Ideally, we would simply go out and see whether our hypothesis fits the facts. But the problem is that

we don't have direct access to the details which would confirm or refute our hypothesis. We don't have access to speakers of Old Icelandic who show or don't show a tendency to have all stressed syllables of the same length or have an irregular tendency to shorten long vowels in front of more than two consonants in stressed syllables. We don't know either exactly what Old Icelandic stress was like. There may also have been other factors at work, like some laws of tonality, which, it is assumed, were the historical origin of the modern Norwegian and Swedish word-tones. There are all sorts of things we would like to know, but we don't. In the absence of relevant data, we can only make guesses which seem more or less likely to be true. Ideally, the likelihood of the truth of hypotheses which are not directly supported by data should be measurable on some sort of an independently justified evaluation scale, based on some notion of natural sound change, which could help us to decide what is most likely to have been the actual state of affairs or course of events. We have already touched on the subject of evaluation of different hypotheses in historical linguistics, and since a further discussion of the matter would lead us too far afield, we will leave it at that and carry on as if we knew what we were talking about.

We will now turn to the question of what caused the lengthening of short vowels. We have already assumed that in Old Icelandic there existed an articulatory unit which we called the stressed syllable. We also assumed that there was a tendency to have these stressed syllables of approxi-

mately the same duration. On these grounds we proposed an explanation of why the old long stressed vowels were shortened before two or more consonants. Obviously, if we believe in these assumptions, we can use them to explain why the old short vowels lengthened. When a syllable like fat 'a piece of clothing', in which a short vowel is followed by only one consonant, was pronounced as a stressed phonetic syllable, we can look at the lengthening of the short vowel as resulting from the tendency to have all the stressed syllables of the same duration. If the vowel was stretched, the duration of the syllable could come to match approximately that of long syllables like fát or fatt 'erect' (neuter).

It should perhaps be emphasized that we are not saying that the vowel was stretched in order for the syllable fat to get the same duration as the long syllables, rather we assume that at the initial stage the lengthening was an automatic consequence of the stressedness of the syllable. This seems to be a reasonable assumption, given the close relation between stress and duration which phonetic studies have shown to exist in many languages. Obviously, it is possible to stretch a short syllable ending in a short vowel followed by one consonant in two ways, that is, either the vowel or the consonant could be stretched. However, there are probably differences in the 'stretchability' of different segments; for example it would seem to be easier to prolong the duration of vowels than that of stops. Other consonant segments, like nasals, would seem to be quite 'stretchable'. At the

earliest stages of the lengthening of old short syllables, it is, then, quite possible that some syllables were lengthened phonetically by stretching the consonant, (as must have been the case in many Swedish and Norwegian dialects), and it is also possible that both segments were sometimes stretched at the same time. Indeed, there are examples to be found which seem to reflect a consonant lengthening. One of these examples is the adverb fram 'forward', which has in Modern Icelandic, when pronounced under stress, a long consonant and a short vowel, but seems to have had a short vowel and a short consonant in Old Icelandic. As we have seen, a lengthening of the consonant is quite common in Norwegian and Swedish dialects, and it has been suggested that lexical doublets in Modern Icelandic, one form showing a long consonant and the other a short one, e.g. ramur:rammur 'strong' (Sigmundsson 1970:325) could derive from a stage of phonetic indeterminacy as to the length of the segments. It seems very likely that at the earliest stages of the phonetic lengthening of old short syllables, there existed forms in which a phonologically short consonant was stretched, but most likely a lengthening of the vowel was more common. When the lengthening of old stressed syllables became a part of the phonological system, the generalization that was made by new speakers was that all stressed vowels were long before one consonant, but in a few cases, like in the form fram, restructuring took place in the lexicon so that the old short consonant was replaced by an underlying geminate.

3.4 A shift or a conspiracy?

In discussing the dating of the changes, and on various other occasions, we have touched on the question of whether the quantity shift was something which we could properly call a 'shift', or whether the term 'conspiracy' should rather be applied. Obviously, in order to be able to deal sensibly with such a question, we would have to give these two notions some clear meaning. One of the things we have assumed could be used to distinguish between the two notions is the timing of the events. If we have reason to believe that some complicated change took place in 'one stage' or over a short period of time, we would presumably tend to call it a shift or a single change, whereas if a number of changes which take place at different times 'conspire' to form a unified, simply storable result (cf. Lass 1974), we would tend to call it a conspiracy. Another factor which will have some bearing on whether we call something a conspiracy or a shift is the cause (or causes) of the change (or changes). If we say that a complicated change was caused by one single factor, we would tend to call it a shift or a single change, whereas if a number of changes which form a unified result (that result being apparently the only thing the changes have in common), are related to different conditions or perhaps no prior conditions at all, the only reason for their taking place perhaps seeming to be the result, we would tend to call that a conspiracy.

We have set forth hypotheses as to what caused the shortening of old long vowels and the lengthening of old short ones in stressed syllables. We related both of these changes to stress, and in that respect we can say that there is a common element in the causes, and consequently we should perhaps call the quantity shift a proper shift. We notice, however, that stress was claimed to be related to the changes in different ways. In the case of the shortening of old long vowels, one can say that the stress was not a direct cause of the shortening, but rather, we assumed that the stress-pattern defined the context in which the shortening took place. We suggested that within the context of an articulatory unit which we called a stressed syllable the long consonantism caused the long vowels to develop shortened allophones. In the case of old short vowels, however, we can say that the stress was a more immediate context for the lengthening, there being assumed a close phonetic connection between stress and duration. In this light we see that the question whether the shortening of old long vowels and the lengthening of old short ones were caused by the same factors cannot be answered simply with a yes or a no. Even though the most immediate phonetic features which we connected the changes with are different, on the one hand a long following consonantism and, on the other, stress and a short following consonantism, we can say that a common factor was working in the background, namely (according to our

assumptions) a tendency to have all stressed syllables of the same duration. One might perhaps make an objection here and say that this common factor is not a historically prior cause as we claim, but simply the conspiratorial result of the two changes. That is, we are perhaps not justified in saying that a tendency for all stressed syllables to have the same duration was a part of the environment of the phonetic processes involved, and all we are allowed to say is that it was a consequence of the two changes that all stressed syllables came to have the same duration (and phonological or prosodic length). If this objection is valid, we could perhaps turn around and say that the same duration of all stressed syllables was the 'aimed-at result' of the changes, to borrow a phrase from Lass (1974:312), and it is not the case that the tendency to obtain the same duration for all syllables was a prior phonetic cause of the changes, but rather that it was the 'purpose' of the changes that all stressed syllables were to have the same duration. In terms of causality, one could then say that the uniformity of the length of stressed syllables was the 'final [orthogenetic] cause' (cf. Lass 1974:312 and 333) of the changes. Looked at in this way the quantity shift seems to qualify as a historical conspiracy.

It seems, then, that the answer to the question whether to call the quantity shift a conspiracy or a shift can depend (apart from the question of the dating)

on whether the tendency for all stressed syllables to be of the same length can be said to have been a relevant factor in the phonetic processes resulting in the systematic changes of vowel shortening and vowel lengthening. We have already suggested that this was the case, but we may of course wonder whether we are right. We can put the question like this: Is it more likely that what the two changes have in common is that they were partly triggered off ^{by} a common environment or that they aimed at the same result, which could then be called the 'final cause' of the change? One might say that the answer to this question will depend on what general view we have of linguistic change and its causes. If we look at language change as in general caused by phonetic and other factors present at the time of change or shortly before it, we would tend to favour the explanation which says that the changes were caused by a phonetic tendency to have all stressed syllables of the same duration. Another way of looking at language change is that it is in general, or sometimes, teleological, that is, the changes that occur sometimes aim at some target reachable in the (near or distant) future. If one accepts that as a more plausible view, one would presumably favour the conspiratorial explanation. Of course the two points of view are not mutually exclusive; it is possible to be an eclectic and say that some changes are teleological and others mechanical, and indeed one can maintain that some (or all) changes are partly teleo-

logical and partly mechanical. 12) I will not embark on a long discussion of these matters, since they involve methodological questions which I am not prepared to deal with. I will content myself with stating that, generally, it seems less to be expected that language change has some predetermined purpose, and therefore, if another explanation is possible, I think that an explanation making that assumption in some way necessary should be avoided. Consequently, I will assume that the lengthening of short vowels and the shortening of long ones are traceable back to a common phonetic element, namely a tendency to have all stressed syllables of the same duration. (The shortening of long vowels, of course, being more indirectly relatable to this tendency as a cause, the more immediate cause being a long consonantism following.)

It should be emphasized that although it seems to me that the term conspiracy should not be applied to refer to the Icelandic quantity shift, that does not mean that it should not be applied to other phenomena. For instance the changes in English discussed by Lass (1974), which gradually, over a long period of time, 'aim at' eliminating vowel quantity as a paradigmatic feature, are much less amenable to a mechanistic explanation of the sort proposed here for the Icelandic quantity shift, because the changes forming the quantity conspiracy in English, culminating in Aitken's law in

Modern Scots are much more formally heterogenous and chronologically far apart than the Icelandic changes seem to have been. It seems that the term 'historical conspiracy' can be used to refer to the English changes, at least until a more satisfactory explanation presents itself which makes the term unnecessary.

Chapter V.

THE PHONOLOGICAL ANALYSIS OF LENGTH.

In the foregoing treatment of quantity in Icelandic and its history, we have carried on without stopping to consider what is meant by the term length or quantity in phonology. In order to try to compensate for this, I will devote the following chapter to the question of what phonological quantity is.

It seems to be general practice to use the terms duration and quantity in such a way that duration denotes absolute physical length, but the term quantity is used to denote duration when it functions 'as an independent variable in the phonological system of a language'. (Lehiste 1970:42) I will try to follow this practice in my discussion. It seems, then, quite convenient to use the term length in any context in which one either does not want to make this distinction or where, for some reason, neither is applicable. The term length can then for example be used when talking about the properties of phonological units when discussed by themselves.

1. Ways of accounting for length

In general, in analyzing length phonologically, there is more than one option open for a linguist. These

alternatives can be summarized (cf. Lehiste 1970:43-44) as being basically four.

It is possible to look upon quantity as being a feature of the segments of the language (most typically vowels, but consonants may also be analysed as long or short). In this analysis, some segments of the language are, then, long, whereas others are short much in the same way as e.g. some vowels are rounded and others unrounded. This feature can be 'distinctive' or 'non-distinctive', just as any other feature can be 'distinctive' or 'non-distinctive'. For quantity to be a 'distinctive feature' of a segment it is a necessary condition that there be at least two degrees of length in the segments, just as roundedness is not distinctive unless there are both rounded and unrounded vowels in the language. There are other conditions for quantity to be phonologically 'distinctive', for example that the length of segments is not predictable by features in the environment. (I will make some further comments on 'distinctiveness' shortly.)

Another way of analyzing length is to assign it more or less a status of phoneme. In this way one can e.g. analyse the difference between Old Icelandic sat and sát as being that in the latter member of the pair there is present a phoneme of length. The pair could be transcribed like this /sat/ vs. /sa:t/, /:/ being the length-phoneme. Obviously this sort of phoneme has a peculiar status compared to other phonemes, for example in that

its phonetic realization varies greatly, being (apart from the element of duration) completely determined by adjacent (preceding ?) segments. This peculiar status, it has been suggested, should be indicated by giving it a special name and calling it 'prosodeme' (Haugen 1949) or 'chroneme' (Jones 1962). This analysis would probably not have very much appeal to others than the most abstract minded (pre-generativist) structuralists. It seems, for example, that there is something missing as far as the representation of the relationship between the length phoneme (chroneme or whatever), and the actual segment that is phonetically long. (See, though, Haugen 1949 who suggests an account of length and other 'prosodemes' (like tones) in terms of timing within the syllable.)

A third way of accounting for length is to assign it to higher order units like syllables, or still larger units, even words (cf. Lehiste 1970:50 ff.) In this sort of analysis, the duration of segments is determined by their position within these larger units, i.e. their relation to other segments and their structural status in the unit. A very simple case of a system describable in this way would be one where vowels are long in open syllables, but short in closed ones. (Thus, every syllable, open or closed, would have similar duration). Questions of 'distinctiveness' may get unclear answers in this sort of situation. This was reflected in our discussion of Modern Icelandic (Chapter II., Section 1.), where there was a question as to which was 'distinctive',

vowel length, consonant length or perhaps something else. A generative analysis of this situation would be one which left segments unmarked as to length at the underlying phonological level and then distributed it on the appropriate segments by a rule defined on the higher order unit (the syllable or whatever). It is, in this connection, an interesting question what sort of an entity length is. Is it appropriate to use the same sort of theoretical notion to represent this and the length that belongs to segments? Can one use a feature [+long] for both, say, a vowel that has been lengthened by a rule of the sort mentioned above and one that is underlyingly long, without making any notational distinction between the two types of length? A similar question arises in connection with palatalization, which can be either a feature of an underlying segment occurring in all (or most, if it is sometimes neutralized) environments or else a secondary feature occurring in special surroundings predictable by rule. It is common to look upon this as different applications of the same feature (cf. e.g. Halle 1959/1971:61-62). If length is treated similarly, one could then say that one and the same feature is in one case inherent to a segment and in another imposed by a phonological rule. The difference is, then, not in the nature of the phonological concept of quantity, (denoted by the feature [+long]), but in its function and place within the system. I am not sure that this is a permissible way of looking at the two different func-

tions of length, I am not even certain that the question of whether or not the two types of quantity functions have a common element is a sensible question to ask. Anyway, what is a phonological feature? Is it something positive that can show up in several languages and different surroundings and be always the same thing, or is it more sensible to look at phonological features as basically relational, having different meanings according to the system that they work in? Certainly, it can be said that a feature like [+high] has a different 'value', in the Saussureian sense, in a system with two vowel heights than when it occurs in a system with three vowel heights. Is there something positively the same about the two instances of the height feature? I am not sure. Their phonetic (acoustic and articulatory) correlates may be similar in some ways, but does that mean that their phonological content (if any) is the same? The basic function of distinctive features is to keep phonemes apart, and their value or content depends on the system they work in.

If the phonological content (value) of a feature [+high] may be different in any two different languages, one may well wonder whether it makes any sense to speak of universal phonological features. It would not surprise me if it turned out that the whole idea of universal distinctive features is wrong, since things like [+high], [+back] etc. can never have exactly the same value in two different languages. I think Trubetzkoy is right when he

says: "Das Phoneminventar einer Sprache ist eigentlich nur ein Korrelat des Systems der phonologischen Oppositionen" and that "in der Phonologie die Hauptrolle ... den distinktiven Oppositionen zukommt" (1958:60). So, from the phonological point of view the thing is not the positive properties of distinctive features, but their negative distinctive function, and systems of oppositions are in principle either wholly different or wholly identical. Parts of these systems can't be identical without the whole systems being identical.

So, if it is in principle so that two different phonological systems have no entities which are exactly the same, then the feature [±long] will have as many values as the systems within which it operates. (Of course, the phonetic correlates may be similar from one language to another, but that does not mean phonological identity.) And a fortiori, the length in a system where it is predictable by rule will be a quite different thing from length in a system where it is phonemic. It may, then, seem grossly misleading to use the same term for both. But granted that we know that length, highness, frontness etc. may never mean exactly the same within grammars of two different languages, we are all right. We don't assume that quantity in Old Icelandic was the same thing as in Modern Icelandic, even though we use the same word to denote them.¹⁾

A fourth way of dealing with length is to analyse 'long' segments as clusters of identical segments. According to that principle the difference between sat and sát in Old Icelandic could be accounted for in such a way that sat would be analysed phonologically as /sat/ with a single vowel, but sát as /saet/ with a double vowel or vowel cluster. There are several things that have been listed as signs of this sort of underlying phonological structure. Trubetzkoy (1958:170-174) lists five phenomena which could be taken as indicating that a language analyses its long vowels as geminates.

If morphological boundaries can fall within a 'long' segment, as in Finnish, for example, where there are paradigms like talo 'house', partitive taloa and kukka 'flower', partitive kukkaa, Trubetzkoy (op.cit. p. 170) suggests, long vowels should be analysed as geminates. The idea is that a + a in this case gives a long vowel, and that this analysis can be extended by analogy to other long vowels of the language. I am not sure that this is a very strong argument, although I may have used it when suggesting an analysis of long consonants in Modern Icelandic as geminates (cf. p. 29) in Chapter II. If, for independent reasons, it seems natural to analyse length of vowels as being, say, an 'inherent feature', it is difficult to see how, when two identical vowels are brought together by morphological rules, they can be analysed phonologically as anything other than long vowels. This argument is based on a sort

of a 'free ride'-principle: since a cluster-analysis seems to be appropriate in this set of cases, one might as well apply it to other cases, even though there is no independent motivation for it in these other cases.

A second indication of geminateness, Trubetzkoy suggests, is to be found when long monophthongs and diphthongs have a similar phonological status. That is, when long vowels and diphthongs seem to form a 'natural class' in that they can be represented as one in some phonological rules. He mentions (pp. 170-171) a case in Central Slovak dialects, where there is a 'rhythmical law', according to which long vowels are shortened after long syllables. Long vowels and diphthongs contribute equally to the length of the syllable that is the environment for the shortening of the following vowel.

A third sign of the geminateness of long vowels is to be found, according to Trubetzkoy, when in phonological processes, long vowels (or syllables) have the same status as two short ones. The archetype of this sort of thing is the stress rule of Classical Latin, which states that stress falls on the second 'mora' preceding the last syllable; if the penultimate syllable was long (i.e. ended in a long vowel or was closed by a consonant), the stress fell there, but if it was short, i.e. ended in a short vowel the stress fell on the antepenultimate syllable.²⁾

Trubetzkoy's two remaining criteria have to do with tonal or accentual variation within long segments. When there is a difference in tone or 'accent' at the beginning and the end of a long vowel, that, Trubetzkoy says, is a sign of 'Zweigliedrigkeit' of the vowel (pp. 172-173). A similar idea is proposed by Woo (1972:24-46), only here the geminateness of vowels is used to justify a particular analysis of the 'moving' tones of Mandarin Chinese. But Woo also mentions some other phenomena which seem to support a geminate-analysis of Chinese long vowels independently of the tone phenomenon. There is a rule according to which in the absence of stress,

- 1) long vowels (VV) become short (V),
- 2) diphthongs (tend to) become monophthongs and
- 3) a sequence of a vowel plus nasal becomes a single vowel. (Woo 1972:35.)

This could perhaps be taken to show that the second 'more' of a long vowel and a segment following a short vowel have a similar status.

Although these criteria may seem plausible enough, it seems to me that one criterion for showing that length is gemination of segments is missing: Let us assume that in some language long vowels are geminates, and let us say that the canonical form for a syllable is CV(V)(C), i.e. the syllable types that occur are CV, CVV, CVVC, CVC, and let us say that its basic vowels are five, a e i o u, and these can combine to form vowel-geminates aa, ii uu, ee, oo. One could say that it is a characteristic

of this language that it allows vowel-segments to combine to form long syllable nuclei. It allows sequences of two identical vowels, and so one would expect it to allow sequences of two nonidentical ones; and granted that this is not prohibited by some other principles, the language should allow any combination of its five segments. (In other words, we might suggest that one of the phonotactic principles of this language is something which says: for every V there is a VV, where V is any vowel of the language.) We would expect the language to have not only vowel clusters (diphthongs) like ai, ei, au, but also oe, iu, ui etc. altogether 20 diphthongs. Of course, one might expect some of these vowel clusters to be excluded by special phonotactic constraints, just as certain consonant clusters are often excluded from syllable initial position, but it seems that these should be looked at as exceptions to a general rule. It might turn out that many of the logically possible vowel clusters could be disposed of as unfit either for articulatory or acoustical reasons, but one might still expect a vowel geminating language to be rich in diphthongs.³⁾

Another thing that would seem to follow from a geminate-analysis of long vowels is connected with this. If long vowels and diphthongs are basically the same, i.e. two short vowels combined, one would expect, other things being equal, that the components of diphthongs

get the same treatment as single short vowels, and each part of the vowel geminates should behave like a single short vowel. This could possibly be detected in historical changes like vowel shifts. If in a system like the one we mentioned above, a is fronted to æ by a context free change, one would expect aa to go to ææ, ai to go to æi and au to go to æu, etc. This criterion, of course, needs a historical perspective which makes it a bit more difficult, but it has been used, successfully, I think, to demonstrate that Old English had a VV structure for its long vowels (Vachek 1959:446). Similarly it is shown by Lass (1976a:94) that some generalizations concerning the great vowel shift in English can be captured easily by analysing long vowels as two identical vowels. For example, changes like that of ee (long e) to ii (long i) and that of eu to iu can be described as a single change of e to i.

To sum up, there are certain things that it seems should go with vowel gemination, and these things can be used as clues as to whether a language has geminates underlying its long vowels. It may be that none of these clues are conclusive, but when one or more of these is present, it can be taken as a sign that long vowels are perhaps underlying geminates.

I have set up a distinction between four ways of analyzing length; as an inherent feature, as a separate phoneme (prosodeme), as predictable or belonging to higher order elements, and as gemination. But it is

conceivable that there exist systems with a sort of a mixture of more than one of these. In Lass and Anderson (1975) and Lass (1974 and 1976a) it is assumed that a geminate analysis can be combined with predictability of length in certain environments. In this analysis a lengthening rule can simply be stated as a gemination rule (and a diphthongization as epenthesis), and conversely, if needed, a shortening rule can be formalized as a deletion. This would seem to be a proper way of doing things, if there is motivation for a geminate analysis of vowels (or consonants) but there are still some synchronic processes at work, changing the length of segments.

2. Types of correspondences between dichotomous systems.

Lass (1976a:45) points out that there are basically two ways (with a cline of mixtures of both) in which two dichotomous phonological systems like the ones split by quantity can be related to each other. It can be a question of corresponding pairs, so that there is a one-to-one function between the two systems, giving pairs where each phoneme from one system corresponds with one from the other. Lass calls this 'pair-based' relation. On the other hand a dichotomous system can be split into two without there being a correspondence between any units in the subsystems. Here, the subsystems are opposed as wholes. All the segments of one subsystem have a

property (or properties?) which members of the other subsystem don't have, but there are no segment-pairs which are kept apart by only the dichotomizing feature. This sort of relation, Lass suggests, should be called 'set-based'. As an example of this type of dichotomy Lass takes the German vowel system, which has two sets of vowels, sometimes said to be distinguished by length, and sometimes by 'tenseness'. Among the things that seem to justify a dichotomy is the fact that only vowels from one of the subsystem (the long or 'tense' ones) can appear in final stressed open syllables. In this type of situation a question to ask is what it is that makes some vowels but not others able to stand at the ends of final open stressed syllables. In principle one can look at this in two ways. One can say that there seems to be some property of the vowels that makes them able to take up those position, i.e. that a phonotactic constraint determining which vowels can stand in this particular position is stated in terms of some feature of the vowels. Or one can say that the ability of the vowels to stand in this position is a property in itself. What the 'tense' vowels have and the 'lax' ones don't is the feature [ability to stand in open final stressed syllables]. An obvious way to try to decide between these two alternatives, is to see if the two sets show similar behaviour in some other respects, to try to see whether some other phonological regularities (or irregularities) can be explained in terms of the 'tense/lax'-dichotomy. In

English, which has a similar distinction, there seems to be justification for a dichotomy independent of the positional argument in that the vowels that can stand in open final stressed syllables show special behaviour with respect to stress-patterns (cf. Lass 1976a:34 and Chomsky and Halle 1968:69 and *passim*). Once this is clear, it becomes desirable to find one property which can be used to explain the two phenomena. It seems to be very clumsy to have two features shared by a set of vowels [ability to stand in open final stressed syllables] and [having function f in stress-rules], or one feature which is these two properties combined into one. But of course the problem still remains of finding one neat property which can be used to account for these two facts (and perhaps some others). Lass has shown that using the feature *tense/lax*, as Chomsky and Halle and many others do, amounts to nothing but inventing a cover-term for the properties in question, since the feature 'tense' seems to be phonetically empty, its phonetic correlates having little or nothing in common but the fact that Chomsky and Halle (as well as others) assign them to the underlying feature 'tense'. So, whatever the motivation for the tenseness feature as a phonological entity, it is not phonetic. Lass suggests that the 'tense' vowels of father, boot, bite etc. be analysed as underlying vowel-clusters. Another line along which it has been suggested that the difference between the two sets of vowels can be accounted for is to emphasize that there is

a difference in the contact the short ('nontense', 'nongeminate') and the long ('tense', 'geminate') vowels make with following consonants (cf. Trubetzkoy 1958:176).

Pairing relations and set-based relations can theoretically be based on any sort of features. One can for example imagine a system where the feature rounded/unrounded splits the system into two subsystems, and there can, theoretically, be a one-to-one function between the two systems, based on the roundness feature. We are interested in cases where there seems to be some plausibility in assigning the dichotomizing function to duration or length. Indeed, it is to be expected that the question of dichotomy in vowel systems arises in connection with length, since length has a very independent status as a 'feature' in vowels. There seems to be very little in the way of markedness relations between length and other features that usually occur in vowels. Whereas we seem to be able to detect some general laws of naturalness (or markedness) of combinations of 'segmental' features like roundedness, frontness and height (high back vowels tend to be rounded etc.), it seems to be much less obvious that quantity as a phonological entity combines more or less favorably with one vocalic feature than others. True, the 'intrinsic duration' of high vowels is seen to be shorter than that of low ones (cf. Lehiste 1970:18-19), but that does not seem to lead to the fact that high vowels are worse adapted than other ones for taking part

in a quantity correlation as a phonological opposition. The fact that length seems to be easily combinable with other vocalic features makes it a good candidate for dichotomizing vowel systems. It seems to be relatively easy to keep everything constant except the duration. That does not mean, of course, that length is always independent. As I have said before, there seems to be, in languages like Icelandic and Faroese, a connection between diphthongal quality and length, and in standard German and Swedish, there seems to be some connection between height and/or peripheralicity and duration (cf. e.g. Lass 1976a:46-49), but it is far from obvious that quantity can be said to be in some 'markedness' relation to other features as for example frontness/backness is to roundness.

Let us now try to clarify the connexion between the question of how to deal with length phonologically and the question of the relation between the subparts of a dichotomous vowel system. It seems that the ideal function of quantity within vowel systems is to establish a paired relation between two isomorphic systems. By adding length to a system, one can get a set of pairs of long and short vowels.

In this sort of situation, where we have a pair-based relation between two systems, it is not a matter of the relation between the subsets themselves how length should be analyzed. It will be determined by consider-

ations of the sort we described above. For example, we may ask whether long vowels behave like two-segment sequences, to see whether we should analyse these as vowel geminates or not. When we look at set-based dichotomies, however, it seems that it is no longer an independent question how to analyze quantity. A priori, it seems that if two long and short subsystems are to some extent in a set-based relation, in that there is not a one-to-one relation between them, the most likely place for length is inside the segments, in other words, that it is an inherent feature.

A set-based relation between a long and short vowel system seems to be theoretically impossible in analyses like the one proposed for Modern Icelandic in Chapter II above, where length is assigned to vowels by a simple rule. The long and short 'systems' (if that term is appropriate here) are automatically related by pairs of long and short (lengthened/shortened (non-lengthened)) vowels. It seems also rather odd to think of length as a prosodeme in the context of dichotomous systems in a set-based relation. A set-based length dichotomy is by definition such that there are no pairs kept apart only by length. So the 'prosodeme' would have to be something more than length. It seems that basically the alternatives that we are left with as plausible ways of analysing length in set-based dichotomies are geminate analysis vs. inherent feature-analysis. In the following I will try to show that a geminate-analysis is basically such

that it only fits a pair-based dichotomy and consequently that in a set-based dichotomy an inherent feature analysis is, other things being equal, the most appropriate one.

If we have a dichotomy apparently based on quantity, but there is not a one-to-one relation between the two sets of vowels, that means that there are some vowels that don't take part in the correlation and/or that some other features go with the length, which make an additional (secondary) distinction between the two systems or potential pairs of vowels. If we wanted to analyse quantity in such a system in terms of gemination, it seems that problems would arise that don't arise in an inherent feature analysis. In the former case, where there are vowels that don't take part in the correlation, we would like to, or should be able to, explain why certain vowels occur as geminates, but not others. If gemination of vowels is to be stated basically in terms of a principle of the sort mentioned above (p.321), it would seem that we would have to explain the non-occurrence of some geminates in terms of special phonotactic constraints. In an analysis giving length the status of an inherent feature this problem does not arise, since there is no reason to expect every vowel to have distinctive length even though some vowels do, any more than roundedness/unroundedness has to be distinctive in every vowel of a system where it occurs. In the second case,

where there are other features that go with the length, the problem is different. Here, we are dealing with a situation where there are no pairs where long and short vowels have the same quality. (It is, of course, theoretically possible that there are some long and short vowels of which it can be said that they have more or less the same quality. But having already assumed that the relation between the subsystems is set-based, we will exclude this possibility for the purposes of our discussion. The fact that there are some pairs that are only kept apart by length could be taken as (some) indication that the subsystems are in a pair-based relation with secondary quality differences in some vowels. It is also possible that one is in cases like these dealing with subsystems in a 'mixed relation', having some pairs and the rest in a set-based relation. I am here, for the sake of the argument, only talking about clear cut cases.)

If we wanted to decide whether a system that on the surface seems to have a set-based dichotomy is analysable in terms of gemination, we would of course go out and look for clues of the sort mentioned above. These clues might lead us to decide that a geminate analysis is indeed appropriate, but I maintain that by that we have automatically claimed (or should have claimed) that the relation between the two systems is fundamentally pair-based. Assuming that there is some logical difference

between a geminate-analysis and an inherent feature-analysis, it seems to me that this should be so as a matter of theory. If a long vowel is analysed as a sequence of two identical vowels, one is in fact claiming that a long vowel is not one unit, but two: [a:] is really not a single long vowel but two segments occurring together. As I have said before, in slightly different words, it seems to me that the fact that a vowel has two parts should lead us to expect that it can be split up and each part can occur independently of the other. So if we have a nucleus like /aa/ we would, granted that the language has single segment (short) nuclei, expect /a/ also to occur, and vice versa. But this would evidently lead us to expect a pair-based system, where there is a set of double syllable nuclei and another set of single nuclei in one-to-one correspondence.

We can put the matter in a slightly different perspective like this: if we have on the surface a dichotomous system, where there are no qualitatively identical pairs of long and short vowels, i.e. the dichotomy seems to be set-based, there are two options open for someone who wants to make an analysis in terms of gemateness. Either one can say that each vowel, long or short, has its own quality, and the difference between the long and the short vowels is that the long ones only occur as geminates and the short ones can only occur as nongeminates. There seems to be no way of capturing this other than by stating for each vowel quality, whether it is geminate or

not. But then of course, one will ask, what is the difference between this and an analysis where length is an inherent feature? Isn't a geminate analysis of this sort equivalent to an inherent feature analysis? If two or more ways are possible for analyzing length phonologically, one would like these to have different empirical (in the Popperian sense) implications. (One would, in other words, like them to be more than 'notational variants' of the same thing). If the prediction of the geminate analysis that the same segment can occur either single or geminated, is taken away from it, as in our hypothetical example, the distinction between geminate analysis and feature-analysis will move closer to emptiness and the grounds for making the distinction weaken. The other way to deal with a surface system that seems to have a set-based dichotomy is to invent an abstract geminate analysis. One can for example suggest to account for two vowels, say [a] and [ɑ:], that both occur in some language, in terms of geminate-ness, setting the long one up as a double and the short one as a single instance of the same vowel by abstracting. This could be done by setting up an abstract entity, which is either underlyingly back (as [ɑ:]) or non-back (as [a]) or perhaps unmarked with respect to backness. The surface forms would then be derived by a rule (or a set of rules) adjusting the backness of the surface segments. But evidently, what is being done by this is to introduce pairing into the abstract system: |a| and |ɑ:|, not being a pair distinguished

only by length on the surface, are made such by analysing them as respectively a single and a double instance of the same phonological element.

I hope that the foregoing has shown that claiming that length, dichotomizing a vowel system into two subsystems, should be analyzed as gemination of vowels implies, or should imply as a matter of principle, that the relation between the two subsystems is fundamentally pair-based. But it has also shown another very important thing, namely the influence that allowing for abstractness has on the way phonological phenomena can be analyzed. If one has a system that on the surface appears to have a set-based dichotomy and seems therefore not suitable for a geminate analysis, it can be made amenable to such an analysis, if a certain amount of abstraction is allowed. Given the possibility of abstraction, all sorts of ways open up for analyzing systems as something other than what they seem to be on the surface. It becomes possible to relate vowels of different surface qualities to the same underlying quality, based on gemination vs. non-gemination, as in the example above. As usual, when abstract analyses are proposed, this should be justified by some regularities that can be captured if the abstract analysis is adopted. For example, if long vowels behave like sequences of two segments with respect to some phonological phenomena, that may be used as an argument for an abstract geminate analysis. But, of course, the abstraction will have to be evaluated as .

involving some cost. This brings out the familiar problem of how to weigh a complication in one part of the grammar against a simplification in another. Deriving a surface quality from a different underlying one, as in our hypothetical example, seems a familiar enough process in generative phonology, and equipped with that, admittedly rather powerful, device one can account for qualitative differences between a pair of vowels differing in their underlying forms only in length (gemination). Systems where, on the surface, not all vowels take part in the long-short relation, as I said before, demand another type of solution in an abstract geminate (by implication pair-based) analysis. Here, it seems most natural to account for the lack of pairing in terms of phonotactic constraints. For example, in a system where all vowel qualities except one appear freely as long or short, it may be plausible to set up a (perhaps independently motivated) phonotactic constraint prohibiting the occurrence of two instances of this particular vowel one after the other within one syllable. In this case the cost of the phonotactic constraint will, of course, have to be measured against the gain made in other parts of the grammar by analyzing long vowels as geminates. It follows from this that the more constraints that have to be set up to prevent the otherwise predicted free occurrence of vowel geminates, the less plausible the geminate analysis becomes, and the greater the other gains made by it will have to be.

From the foregoing it should be clear that allowing for abstractness in phonological solutions, leaving aside the question of the linguistic plausibility of abstract solutions in general, complicates the distinction we are discussing between geminate-ness or 'inherentness' of phonological length. Long and short systems that on the surface seem to be in a set-based relation and therefore suited for an inherent feature-analysis can be, by abstraction, made amenable to a geminate analysis, which, I think, implies a pair-based relation between the two subsystems.⁴⁾

3. Diphthongs and long vowels

Up to now I have paid little attention to the relation between diphthongs and long vowels. I suggested that, other things being equal, a vowel-geminating language should be expected to be relatively rich in diphthongs, its basic principle being that vowels could combine freely to form vocalic nuclei of double length. I also mentioned that Trubetzkoy suggested that if long vowels and diphthongs behave similarly, that could be taken as a sign that the language treats its long vowels as geminates.

Granted that it has been established that a language analyses long vowels as geminates, it seems to follow that its diphthongs should be analysed as vowel clusters. (This is of course circular to the extent that Trubetzkoy's diphthong-criterion just mentioned is

used to establish the geminate-ness of the long vowels, but we will ignore that for the moment.) But if it has been decided on independent grounds that vowel length in a language does not derive from geminate-ness, rather that it is a feature (either inherent or assigned by rule) of the vowels that is responsible for the long duration, it is a different question how to deal with diphthongs.

We can say that we are really faced with the question of what a diphthong is. (From the phonetic point of view, one can say that a diphthong is a vocalic part of one syllable that ends with a different quality from the one it started with, but from ^{the} phonological point of view something more is needed.) According to my general belief that phonological entities are largely language-specific, I shall maintain that diphthongs can be different things in different languages. If diphthongs are to be analysed as vowel-clusters in languages that treat their long vowels as geminates, it seems that in languages that have length as a feature of segments (at some level, segmental or derived by rule), diphthongs are most typically simply vowels with two qualitatively different stages, moving vowels, if you like. This corresponds in part at least with H. Andersen's (1972:18) distinction between 'sequential diphthongs' and 'segmental diphthongs'. 'Segmental diphthongs', according to Andersen's definition, are single segments 'whose central phase is acoustically heterogeneous in its temporal development'. A 'sequential

diphthong, on the other hand, is a 'sequence of segments, usually forming part of the same syllable'. I will use these terms in the following, although I am not sure that I use them in exactly the same sense as Andersen. (It seems to me that some comments of Andersen's concerning 'phonetic and phonological diphthongization' may indicate that he considers that 'segmental diphthongs' can not be phonemically (underlyingly) defined as such. I would not want to subscribe to this.)

It should be made clear that I am not claiming that there is an if-and-only-if relation between length as a segmental feature and 'segmental diphthongs' or between length as gemination and 'sequential diphthongs'. I am only saying that given one thing, the other is to be expected. I think that, in general, linguistic phenomena are such that 'if-and-only-if' statements, involving the universal quantifier, are impossible.

A third type of diphthongs should perhaps be added to the two types mentioned above. I have in mind diphthongs that arise in morphological alternations when more or less vocalic segments, that there is reason to analyse as monophthongal segments in the underlying system, combine to form phonetic stretches that have the sort of acoustic and articulatory patterns that characterize diphthongs. These could perhaps be called 'combinatorial diphthongs'. An example of this type of thing could be the Modern Icelandic (except for south-eastern dialects) alternation between [ɔ:] and [ɔi], [Y:] and [Yi], and [I:] and [Ii]

in forms like bogi [bɔ̃jI] vs. boga [bɔ̃:ɣa] 'bow' (nom. vs. acc.), Hugi [hYjI] vs. Huga [hY:ɣa] (a man's name) (nom. vs. acc.) and stigi [stIjI] or [stij:I] vs. stiga [stI:ɣa] 'ladder' (nom. vs. acc.). In the nominative forms of these words the voiced fricative following the vowel is palatalized by the following front vowel, but in the accusative it appears as a velar in front of a back vowel. When the palatal fricative meets the preceding vowel, a sort of a high glide appears, 'connecting' the vowel with the fricative. There seems to be good reason to analyse the morphemes in question as having underlying monophthongs and to account for the alternation accordingly, for example because the 'diphthongs' [ɔ̃j], [Ij] and [Yj] don't occur in other environments, and [ɣ] does not occur before a front vowel. We seem to have, then, a case where features deriving from segments that follow the vowel 'move into it' to make it diphthongal. Similarly in English, the diphthong in forms like day derives historically from a sequence of ɔ̃ + a palatal fricative, and while this process still was synchronically active, one could perhaps say that the form had a sort of a 'combinatorial diphthong' resulting from the concatenation of ɔ̃ and j. It will of course have to be justified in each case that these 'combinatorial diphthongs' have a different phonological status from other, more 'deeply rooted' diphthongs.

An important question that I will leave undiscussed here is the question of the 'syllabicity' of the two

parts of diphthongs. I will, in other words, not discuss the distinction between a 'rising' and a 'falling' diphthong, which is commonly made. I hope this does not invalidate the other comments I am making about the nature of diphthongs.

If it is true that there are 'segmental diphthongs' with an internal movement of quality, it must follow that it is not necessary that phonological features like [high], [front] and [rounded] always have whole segments as their domain. One can then have segments that are [-high] (or mid) at the beginning but [+high] at the end: [ei], or rounded at the beginning but unrounded at the end: [ɔa] etc. Although this fits rather badly with the general practice in using phonological features (cf. Chomsky and Halle 1968 and Ladefoged 1971), it seems to be inevitable to allow for such 'movements' in phonological quality within single segments. This sort of thing is by no means confined to vocalic diphthongs. There occur in languages in many parts of the world sounds that may be called consonantal diphthongs, for example pre- and post-nasalized consonants [ṁb, ṁd, ṁg, ḃm] etc. (cf. S. Anderson 1976). These sounds behave phonologically as single segments, but have a complex articulation, beginning as nasals, but ending as stops or vice versa. Anderson suggests that these, and nasals in general, should be treated as "oral stops on which a nasal pattern is realized: if the stop is nasal throughout, we get the common primary nasals, while 'contour' nasality patterns

give rise to pre- and post-nasalized stops." (op. cit.: 343). Anderson also considers the possibility of treating affricates and labiovelars in a similar manner, but is rather in doubt that this is appropriate. Anyway, the fact that both pre- and post-nasalized consonants occur and behave like single segments, seems to support the idea that diphthongs can be treated as single segments with changing articulation. What sort of formalism is needed for this is a different question which I will not discuss here.

4. The phonological analysis of length in Old Icelandic.

Having given a rough outline of some of the points that arise in general in connexion with questions of how to analyse length phonologically, I will now move on to consider what sort of analyses are appropriate for the two stages of Icelandic that we have been concerned with in the earlier chapters of this study. I will start with Old Icelandic and then have a second look at the analysis proposed in Chapter II for Modern Icelandic.

4.1 The nature of length about 1200

To summarize briefly what was said about the vowel system of Old Icelandic in Section 1.1 of Chapter IV, the vowel system of Icelandic shortly after 1200 probably looked something like this:

short			long			diphthongs		
i	y	u	i:	y:	u:	ei	ey	au
e	ø	o	e:	ø:	o:			
a	ǫ		æ:		ɑ:			

Later in the 13th century, the merger of short /ø/ and /ǫ/ into one phoneme, designated /ö/ and of /ø:/ and /æ:/ into something represented as /æ:/ the system had become something like this:

short			long			diphthongs		
i	y	u	i:	y:	u:	ei	ey	au
e	ö	o	e:		o:			
a			æ:		ɑ:			

The diphthongs at both these stages had the same 'prosodic' status as long vowels. This can be deduced from the fact that they, like the long vowels, in concatenation with a single following consonant formed sequences that could function as monosyllabic ictuses in poetry.

If we now try to apply the notions we described above to these data, we can ask ourselves what sort of an analysis is appropriate for length in Old Icelandic. Should we analyse it as, (1) an intrinsic feature of the vowels, (2) gemination of vowels, (3) a separate phoneme, or (4) belonging to a higher order element, say a syllable?

It should be relatively safe to exclude the last two alternatives. The last alternative seems to be out since there are minimal pairs distinguished by long and short vowels: fat 'a piece of clothing': fát 'confusion':

lit 'colour': lit 'I look', vel 'well': vel 'a trick' etc. It seems to make no sense to set up, say, different syllable types to distinguish between these pairs only to have this difference appear as a surface distinction between the vowels. The phoneme (prosodeme) analysis seems implausible on general theoretical grounds and because the diphthongs have a long duration, which it would be unnatural to assign to an underlying length phoneme, since there are no short diphthongs.

We are, then, left with an inherent feature-analysis or an analysis in terms of gemination. If we start by trying to discover the relationship between the two systems, we discover that it is not entirely clear whether it is set-based or pair-based. As far as limits on distribution are concerned, there seem to be no constraints except that, according to one theory, only long vowels and diphthongs can appear when no consonant follows (in stressed final open syllables and in front of hiatus). There are no forms like, *bu or (bisyllabic) *bua, only forms represented in the standardized orthography as bú and búa. According to Benediktsson (1968) vowel length is neutralized in these environments, and the phonetic duration varies. From a generative point of view, one would of course have to decide what to put in the underlying forms, long or short, and it seems that comparative evidence favours long vowels as the underlying segments in these positions: In Modern Icelandic the vowels of búa and bú have the MI reflex of OI long /u:/ and in most cases these

vowels can be traced back to Proto-Germanic long vowels. So, it is possible that we have here a phonotactic constraint based on the long-short distinction. The diphthongs show the same behaviour as long vowels, i.e. they can occur in stressed syllables without a following consonant, as in hey 'hay'. This, then, could be used as an argument for distinguishing between two sets of vowels in Old Icelandic: the short vowels vs. the long ones and the diphthongs. We have already mentioned the difference in metrical function of these two sets of vowels. There again, we had the long vowels and the diphthongs forming a class as opposed to the short monophthongs. Looking at these facts we seem to have a good case for setting up a dichotomous system, but we may seem to have a rather poor candidate for a pair-based relation. True, we have (around 1200) correspondances like /i/ - /i:/, /y/ - /y:/, /u/ - /u:/, /e/ - /e:/, /o/ - /o:/, /ø/ - /ø:/, but that is about it, except if we want to say that /a/ and /ɑ:/ constitute a pair. There is no long phoneme corresponding to short /a/ and no short one corresponding to /æ:/, and there are no short diphthongs. (The lack of short diphthongs would, of course, not be strange in a geminate-account of the system, the diphthongs being vowel clusters. I will come to this shortly.) Another thing that will have a bearing on the sort of relation that holds between the two systems is the amount of qualitative similarity between the members of the potential pairs /i/ - /i:/,

/y/ - /y:/, /u/ - /u:/, /e/ - /e:/, /o/ - /o:/ and /ø/ - /ø:/. There is evidence that already in the 13th century the corresponding long and short vowels were not completely identical in quality. This is shown by a change that took place in the 13th century in the spelling of the unstressed vowels /I/ and /U/, which, having been spelled e and o respectively, came to be identified with short /i/ and /u/ and spelled accordingly as i and u. This, according to Benediktsson (1962, cf. 1965:72-73), was caused by a lowering in the short stressed vowels /i/ and /u/, that brought them closer to corresponding in quality with the unstressed vowels. It is further likely that the long vowels, /e:/, /æ:/ and /a:/, and perhaps /o:/ as well, had started to diphthongize in the 13th century (cf. above pp.158-61). Both of these changes must have diminished the degree of 'paired-ness' of the correspondence between the two subsystems in that the difference between the members of the pairs /i/ - /i:/, /u/ - /u:/, /e/ - /e:/ and /o/ - /o:/ was now not based on duration alone. This discrepancy between the long and the short system was further increased in the 13th century by the merger of /ø/ and /q/ and /æ:/ and /ø:/. There was now no long segment corresponding to /ö/ (< /ø/, /q/), and the new /æ:/ (< /æ:/, /ø:/) had no short correspondent.

Evidently, the answer to the question of what sort of correspondence prevailed between the long and short systems will depend on which stage we are talking about.

The farther down the time scale we move, the more the relation comes to look like being set-based. And if what I said before about the connexion between the sort of relation that holds between two subsystems of a dichotomous vowel system and the analysis of length in terms of geminateness or as an inherent feature is correct, then the closer we get to the stage when the quantity shift started to have its effects, the less attractive becomes the geminate analysis of length.

It seems, however, that around 1200 the case for a geminate analysis is not so bad. As I mentioned before, the place of diphthongs in the dichotomy, having the same status as long vowels, is natural in a geminate analysis. The diphthongs would obviously be analysed as vowel clusters and should therefore have a phonological status similar to the long vowels. Further, there is a considerable degree of correspondence between the long and short vowels, and the facts mentioned above as speaking against a pair-based geminate analysis are mostly later than 1200. Granted that we are willing to derive /a/ and /ɑ:/ from the same abstract quality, by either a backing rule for the long variant or a fronting rule for the short variant or a two-sided rule, backing the long variant and making the short one non-back, the main problems are the lack of a short version of /ɑ:/, and the lack of a long version of /ə/. But there are good reasons to believe that at slightly earlier stages, both of these were present.

It used to be a common opinion that earliest Old Icelandic had a low front short vowel, the so called 'Umlauts-e' (derived from a by i-umlaut), often denoted ę, distinct from original e. Benediktsson (1964:101) concludes that a distinction between e (< Germanic e, i) and ę (in our notation /æ/) (<a) did not exist around 1200. But there are to be found in the very oldest manuscripts signs of a distinction in spelling between two e-sounds, which, according to Benediktsson, could be taken as indicating a distinction made in earlier (now lost) manuscripts and simply copied by the younger writers. So it is possible that around 1100 there still was a short counterpart to the long e. And if we assume that around 1100 the long and short subsystems (and the nasalized system postulated by the First Grammarian, cf. Benediktsson 1972:130-137) were in a pair-based correlation, one can perhaps expect that this relation would survive a minor blow like the one of the merger of /e/ and /æ/, leaving a gap in the short subsystem. This gap could be, from the synchronic point of view, just an idiosyncrasy caused by a historical change, foreign to the pairing principle which still was valid in general.

Concerning the lack of a long counterpart to the short /ę/, it is even probable that in the first decades of the 13th century there still prevailed in some dialects a distinction between a long /o: / and /a: /. The former had arisen historically as a u-umlaut variant of long /a: /, just as /ę/ was created by the umlauting of short

/a/, but in the early 13th century this long /q:/ disappeared, merging with /o:/ in many nasal environments (nótt 'night' < nōtt < Gmc. *nahtu (the Gmc. a is lengthened before the disappearing h)), but with the original /a:/ in other environments (rákum 'we drove' rókum, < Proto-Scandinavian *rakumR) (cf. Benediktsson 1965:61-62).

So, granted that the distinction between /q:/ and /a:/ prevailed up to about 1200 and that the merger of /e/ and /æ/ was not earlier than about 1100, it is not at all implausible to assume that the long and the short vowel systems around 1200 were fundamentally in a pair-based relation even though they had been and were being hit by some disruptive changes, bringing about some holes in the pattern. If the relation between the long and short monophthongs was pair-based and the diphthongs behaved like long vowels, it seems natural to assume that vowel length in Old Icelandic up to about 1200 was gemination (or cluster formation, in the case of diphthongs) of vowels.

One final point should be touched on in this connexion, and that is the number of diphthongs. In the vowel systems cited above, there are only listed three diphthongs: ei, ey, and au. I have mentioned that languages whose long vowels are geminates should have a tendency to be rich in diphthongs. In the light of what is said above it may seem rather little for a language with eight or more vowel qualities and a principle of vowel gemination and cluster

forming to have only three diphthongs. In response to this criticism, it might be said in the first place, that these are not the only diphthongs that can or have been set up for the earliest stages of Icelandic. There are other sequences that have been analyzed as 'rising diphthongs', beginning in a semi-syllabic i, usually denoted in the standardized orthography by i or j. These are iu (or iú) io (ió) (both deriving from IE eu), ia (ia) and io (io) (arising from 'breaking' of e), and perhaps a fifth ia: (iá) (distinct from ia/ia.) In addition to this, some comments made by the First Grammarian seem to suggest that he looked upon the OI reflex of the MI labiodental fricative [v], when preceding a vowel, as a nonsyllabic instance of /u/, since in listing examples of cases where a vowel 'gives up its nature and must then be called a consonant rather than a vowel' he cites the example vin, 'wine' (MI vin [vi:n]). By analogy with this, one should then expect the First Grammarian to consider any combination of y (perhaps phonetically w) with a following vowel as a rising diphthong (i.e. he would consider y preceding a vowel to be a non-syllabic /u/). However, in other instances he seems to treat it as a consonant, so there is some ambiguity here (cf. Benediktsson 1972:154-155).

In general, it seems that it is doubtful that these 'rising diphthongs' can be used to argue for a geminate analysis of length. It seems that, if they were diphthongs at all, their status was peculiar, not only in

that they were 'rising' but also in that there seems to have been a distinction between 'long' and 'short' diphthongs of this type. Forms with iu (iú), io (ió) and ia: (iá) as their stressed nuclei function metrically as long in poetry, but the 'breaking' diphthongs ia and io function as short. The idea of short rising diphthongs /ia/ and /io/ as opposed to long rising diphthongs /iu:/, /io:/ and /ia:/, seems not to fit at all into the geminate length model. The only way a cluster of i + a vowel can form a short nucleus in a geminate analysis is if the i functions as a consonant, and becomes a part of the onset of the syllable. Similarly, the 'long rising diphthongs' would, within a vowel clustering-framework, be underlying 'triphthongs' /iuu/, /ioo/ and /iaa/, and would only conform to the canonical syllable structure (not allowing, in general, vocalic nuclei of more than two moræ), by either shortening the second quality (by degemination) or by making the i consonantal and assigning it to the consonantal onset of the syllable. These problems would not arise if the 'rising diphthongs' were simply analyzable as clusters of a consonantal /j/ + vowel, as seems to be appropriate for Modern Icelandic (cf. above Chapter IV, 2.3). This is, however, not without problems, since in poetry all the way down to the 16th century, (initial) i alliterates with vowels. (Þórólfsson 1925:XXV-XXVI)). In general, the whole problem of how to analyze the 'rising diphthongs' and the question of the phonological status of prevocalic i (j) and u (y)

looks rather complicated historically and there is not room here for a detailed and sensible discussion of it.⁵⁾

Even though it seems difficult to get support for the geminate-analysis of length from the 'rising diphthongs', that and/or the paucity of diphthongs does not ruin the case for it. We can still make a distinction that will help us to ignore this point. I suggested above that it could be said of a language that is geminating in that it allows every vowel to occur either single or double and, ideally, should allow clusters of dissimilar vowels to form a relatively great number of diphthongs, that a part of its phonotactic principles was something like:

For every V there is a VV.

(This is supposed to mean that every vowel can combine with another vowel (as well as itself) to form a long vocalic nucleus.) If such a system were to exist (I don't know that it does; perhaps Finnish is close, cf. fn. 3), it could be said that the rule above would generate all the long/short vowel pairs of the language, as well as the diphthongs. Suppose a new vowel were to be added into this system; then the phonotactic principle above should allow for it to occur both as long and as short. (It is of course a different matter whether lexical items can be supplied to fill the spaces allowed for; there might be 'accidental gaps'.) Suppose, on the other hand, that things were to happen to this system so that it would not be possible to apply the geminating principle in all cases. If for example there is a merger of two vowels x and y into z in the short subsystem, leaving xx and yy

in the long one without counterparts, and creating z without a corresponding zz in the long system, this would mean that the principle stated above would not hold for all vowels. But it could still be valid for the rest of the vowels in the system and be kept with modifications as a matter of inheritance. One could say in this case that the vowel system (looking back) is still geminating, but other forces have undermined the effect of the geminating rule; it can be said to be 'passive', or 'past-oriented' (cf. Anttila 1975).

If Old Icelandic about 1200 is to be analyzed as having vowel length which is basically vowel-clustering (and gemination), I would maintain that it was so as a matter of inheritance, and if we allow for that, the lack of symmetry between the long and short subsystems and the relative paucity of diphthongs can be looked on as a consequence of the 'passivity' of the gemination. When the umlauts produced new vowels, they did not have to combine with the old vowels and other umlaut sounds to form new diphthongs, and when mergers took place in one of the subsystems, reducing the number of vowels, that did not have to affect the other one. The geminate-ness or clustering relation between the two subsystems could still hold within old established pairs although new vowels appeared that did not take part in it, and old ones disappeared leaving some vowels without correspondents. Thus the geminate-ness of vowel length (if it

existed) was a matter of inheritance that still prevailed in spite of some unfavorable events. Evidently, the sort of things that happened to the old system were slowly wiping out the signs of the old pair-based relation and the geminate-ness of vowel length.

This is exactly what I will assume, i.e. that the geminate-ness of vowel length that perhaps prevailed as far down as to around 1200, gave way to an inherent feature of length in the 13th century. After the mergers and qualitative changes that we have described in this chapter and Chapter IV, it becomes more and more difficult to maintain that the systems are in a pair-based relation, and, proportionally, it becomes less likely that the length is geminate-ness.

Once the geminate-ness was given up and the inherent feature [\pm long] introduced instead, it is obvious that length had a completely different status within the system from what it had before. If we imagine that length was an inherent property of some vowels, but not others, there is no strong reason to expect there to be pairs that are identical in everything but length. On the side of the formalism, a feature like [\pm long] has a similar status to [\pm rounded], and there is no necessity for it to have a minimal distinctive function in every instance. There is now less reason to expect that length is superimposed on other features and that vowels that differ in length only differ in that respect.

4.2 The status of length as an inherent feature

I am gradually getting closer to a problem that I have touched on before (Chapter IV, Section 1.2.2), namely the question of the status of length relative to other features in the Icelandic vowel system from the 13th century down to the time when the quantity shift was completed. Having assumed that length had become an inherent feature of vowels the question arises what sort of a relationship prevailed between the length feature (responsible for the long duration of the appropriate vowels) and other vowel features. We may wonder whether some features of the long vowels or, conversely, some features of the short vowels are predictable from length vs. shortness. Or, is it perhaps the case that length is predictable from some other feature(s)? Should one perhaps invent an abstract feature that can be used to predict both the length and some other features of the vowels (cf. Sigmundsson 1970)? Indeed, do we have to worry about this at all? Is length perhaps just a feature like any other with no special relation to other features? Is, for example, the relation between the positive value of length and diphthongal quality that we have seen probably prevailed quite early in the vowels /æ:/, /ɑ:/, /e:/ and /o:/, to be accounted for in the same way as the relation between roundness, height and frontness? An important point here is whether there really is any need for abstracting and setting up some hierarchical order of the features. Several things may have a bearing on

this. To start with, the system is dichotomous. There are two facts that show this: (1) Long vowels form long 'syllables' when combined with a single following consonant; these have a specific function in metrics. (2) Only long vowels and diphthongs can appear (we assume) in a stressed position when no consonant follows. It seems that neither of these facts demands an abstraction of any sort or a statement of the hierarchy of features. The length just is there and can be used in accounting for this. A more important question is perhaps whether some features are secondary, or redundant. One can say, for example, that the diphthongal quality of the long vowels just mentioned is 'non-distinctive' if we have already distinguished the vowels in question from every other vowel in the system, /æ:/ as a long, front, low, spread vowel, /ɑ:/ as a long, back, low (perhaps rounded) vowel, /e:/ as a long front mid spread vowel and /o:/ as a long back mid rounded vowel. But of course one can ask: Why not have it the other way around and 'predict' the length of these vowels from their diphthongal character. Or is there any reason to predict anything? There are two types of arguments, it seems to me, that could be put forth for an abstraction by which some features are in some sense secondary and predictable from some other underlying feature(s). There are considerations of formal simplicity of the sort just mentioned, that is that one should use the fewest possible features to distinguish between all

the vowels of the system and make the others predictable by rule. Although this type of argument is very widely used in modern linguistics, it is, I think, a bit dangerous (cf. Arnason forthcoming) for very important reasons. It is not, for example, clear what justification there is for assuming that simplicity in terms of an invented formalism can be used to justify a linguistic analysis. In other words, there seems to me to be no guarantee that the laws of the formalism of distinctive features and redundancy rules, which would be used in our case, can be used to make 'generalizations' that are linguistically significant (cf. Lass 1976b). Another type of argument of a similar sort sounds more linguistic. This is that some of the features actually characterizing the vowels have a more central function than the other ones. They are distinctive, the others are non-distinctive. I have already touched on the ambiguity of the term distinctive (Chapter II, Section 1). I suggested that as a technical term in structural phonology, it should only be used in a sense something like 'not predictable from other phonological features'. This functional argument is then similar in its effects to the formal argument above; it leads to a distinction between predictable and non-predictable features, only here it is assumed that the motivation for the abstraction is linguistic. It is not notational economy that is appealed to, rather it is claimed that the linguistic system is such that one feature is basic and the others are derived from it. This

would be reflected in the fact that speakers 'see' the phonemes in question in our case /e:/, /ɛ:/, /o:/ and /ɑ:/, as basically long but having diphthongal qualities as secondary features (cf. Durand 1939). This sort of statement is basically a claim about facts, and is either true or not true. The argument, or perhaps rather the claim, seems to me to be purely linguistic and, in theory, independent of formal considerations, but the problem with it is of course the difficulty in verifying it (our old problem from Chapter IV, §3.1).

The same type of arguments can be set forth for an analysis in terms of 'tenseness' as an abstract feature distinguishing between the long vowels and diphthongs on the one hand and the short vowels on the other. One may argue that it is formally simpler (true or false) to do this, or one may claim that in reality speakers 'saw' the two sets of vowels as different, but connected the difference with no specific phonetic characteristic, neither long duration nor diphthongal quality or anything else, but added these together and subsumed them under the term 'tense', or as is the common practice in contemporary Icelandic school-books, 'broad'.

I have to admit that I see very little evidence that could be adduced in choosing between these alternatives. This is so because the possible difference in 'empirical predictions' (in a Popperian sense) made by the different alternatives are difficult to test from the synchronic

point of view. If one were to say, for example, that length was basic, one would perhaps expect the secondary features (like the diphthongal quality of the non-high long 'monophthongs') to fluctuate. But since we are dealing with a historical ('ideal') stage, and there are no signs of a fluctuation of this sort in the spelling in manuscripts we can't test this. The only evidence we can produce is comparative, and this shows that at a later stage the qualitative differences between the vowels took over the function previously (perhaps) held by the length. But this was a consequence of the quantity shift which we have already assumed to have arisen from causes that, strictly speaking, lie outside the context of relations between segmental phonological features. We can, of course, say that the lengthening and shortening processes eventually undermined the distinctive function of the length feature, and once these processes began to have their effect, the length feature could no longer be central, but that does not tell us anything about the status of the length feature relative to other features before these changes began to have a serious effect on the system.

In spite of this uncertainty we can set up the following as a plausible model of what happened: In the 13th century at the latest the underlying nature of the length correlation changed from being that of gemination and cluster-formation (which was an inherited characteristic) to being a matter of an inherent feature of vocalic seg-

ments. In the period immediately after this, it is likely that the length was a 'central' feature in the monophthongs (having perhaps a different status in the diphthongs), but gradually, as a consequence of the lengthening and shortening processes that later resulted in the quantity shift, the length feature lost its central status and the qualitative features assumed the distinctive role previously held by length.

5. The phonological analysis of length in Modern Icelandic, a quick second look.

Having dealt relatively thoroughly with length in Modern Icelandic in Chapter II, I will make only a few comments concerning this here. My conclusion was that length is predictable in vowels on the basis of stress and the consonantism following within the same syllable.

To put this in the perspective of the present chapter we can say that the relation between long and short vowels is pair-based; every vowel, diphthongal or monophthongal, has both a long and a short alternative, depending on the environment. We can also say that length does not belong on the segmental level, but is a feature derivable from the environment, most abstractly belonging to the stressed syllable. It is limited in its distribution in that it only appears on the vowels (perhaps as 'tenseness' or 'half-length' of consonants between a short vowel and another consonant as in hestur [hes·tYr] 'horse'; cf. Chapter

II, Section 3.3) (It seems natural that length is more or less limited to vowels, since vowels are the most 'stretchable' segments, having a maximally open articulation and greater resonance than consonants.) In terms of the nature of length and the appropriate way to analyze it, it is impossible to say that the length rule can be formulated in terms of gemination, since long diphthongs are not geminated short diphthongs, nor are short diphthongs produced by deleting one mora of a 'full' (long?) vowel. What happens is that the segments, diphthongal or monophthongal, are stretched or shortened as wholes.⁶⁾ It seems, thus, most plausible to look upon length in Modern Icelandic as a derived feature of the vowels (we can call it [+long]) which indicates that the segment in question is relatively long in duration. It must be noted, incidentally, that the fact that diphthongs can appear either as long or short is strong evidence in favor of our suggestion above that diphthongs can be segments with movable quality, but are not necessarily clusters of vowels (cf. Lass 1974, fn.9, p. 339).

One point remains to be touched on concerning length in Modern Icelandic. It has been showed by Garnes (1974b) that there are, in spite of what we have said so far, slight differences in phonetic quality between the corresponding long and short allophones of vowels. For example the high vowels /i/ and /u/ are more diffuse when long than when short, and the diphthongs, /ei/, /öy/, /ai/, /au/, and /ou/ are 'somewhat monophthongized' when short

(Garnes 1974b:3). Perhaps the most important of these qualitative alternations is the slight diphthongal quality of the long variants of the non-high, non-low monophthongs /ɛ/, /œ/ and /ɔ/. Garnes (Table 1.) transcribes the long alternatives respectively as [e^ɛ:], [øœ:] and [oɔ:], but the corresponding short ones as [ɛ], [œ] and [ɔ]. So, from the point of view of the short vowels, the long alternants have raised their first part, thus forming a sort of diphthong. None of this is very surprising, and no doubt the diphthongal nature of the long variants of the mid monophthongs has a quite different phonological status from the diphthongal quality of the underlying diphthongs. It is only to be expected that vowels with a long duration are susceptible to slight variation in their quality, and we can see here the seeds of a diphthongization similar to the one that seems to have hit the long variant of old /a/ in Faroese (see Chapter III, Section 1.).

Evidently, these variations of the quality between the long and short allophones of the vowels are at a low level in the phonology. But it would be a mistake to shrug them off as 'purely phonetic' and therefore uninteresting from the linguistic point of view. The mere fact that these alternations are regular makes them, in some sense at least, a part of the linguistic system, and they cannot be said to be governed by general laws of a purely phonetic sort. But it seems difficult to say what their status is relative on the one hand to firmly established

phonological processes like the length rule in Chapter II and on the other hand to more or less accidental or uncontrollable features of coarticulation etc. Indeed, there is probably one more type of processes existing between the highly systematic level on which rules like the length rule and (morpho-) phonological processes like devoicing of /r/ in front of a voiceless obstruent (vor [vɔ:r] vors [vɔ:rs] 'spring' (nom. vs. gen.) exist and uncontrolled processes of coarticulation etc.. I am talking about things that occur in 'allegro speech', that are shown (e.g. Zwicky 1972 and Dressler 1975.) to be regular and dialect specific to a certain extent. These things probably abound in every language. In Icelandic, we find alternative pronunciations, like enskur 'English' (cf. above p.64): [ɛn·skYr], [ɛskYr] and clines like in vinfengi 'friendship' (a compound word): [vI:nfɛipɕI], [vImfɛipɕI], [vIvɛipɕI] and, still worse, complete loss of stressed vowels as in viltu (koma) 'will you (come)': [vI+tY], [vɿtY], [ɿtY] (with a syllabic (in some sense) lateral fricative in the last two cases). And compare these four versions of það getur verið 'it may be': [θa:ð ʒɛ:t^hYr vɛɛ:rIð], [θaðʒɛ:t^hYr vɛɛ:rIð], [θa:ʒet^hYrvɛɛ:rIð], [haʒ^ɿt^hYrvɛɛ:rIð]. Only the last of these alternatives can be said to be really fast speech. (All of these pronunciations are my own.) It seems, a priori, that these processes in Icelandic cannot be said to be wholly unsystematic and linguistically uninteresting. Already from these four

examples we see a very peculiar regularity, namely that reduction seems to be more liable to take place in syllables that are stressed rather than unstressed on the word level. We get reduction in the nasal following the stressed vowel in enskur, a similar thing in vinfengi, and in viltu and getur (of það getur verið), the vowels that appear in the stressed position in the word are affected more seriously than the 'unstressed' ones in faster speech, they can be literally wiped out (!). This sort of thing, it seems to me, is difficult to explain as some sort of consequences of general phonetic laws of coarticulation etc.

It remains (for me at least) to be seen whether there is a genuine difference of kind between the three types of rules we have seen, the well established length rule, the perhaps not so well established, or secondary, rule of quality variation between long and short vowels, and the reduction rules of fast speech (which, indeed, only reduce certain phonological forms and leave others more or less intact). If these are three different kinds of rules, it would seem to follow that there is much more complicated stratification in phonology than is usually implied in the literature.

Evidently this question of stratification in phonology is too important and complex to be dealt with sensibly in the context of this study, so I will leave it. But to return briefly to the relation of the rules of quali-

tative differences in the vowels to the length correlation itself, let me say this: It seems evident that the quality-variation is dependent on the length variation, at least historically, and it seems to make little sense to imagine anything else from the synchronic point of view. In terms of rules and rule-ordering relations, the length rule feeds the rules for quality alternation. But does that mean that the quality rules are of a kind different from the length rule (and the stress rule, and the syllabification rule proposed in Chapter II)? That is not clear at all. Indeed, one could imagine all of these processes as one big block of intrinsically ordered rules. The stress rule and the syllabification principle define the input of the length rule, which in turn defines the input of the rules accounting for the quality alternations. If that were the case, one would presumably say that all of these rules belong to the same part of the grammar. But, still, one might say that the quality alternation rules are somewhat more 'phoneticky' than the length rule. Is it then the case that all phonological rules are basically the same, but that there is a cline of 'phoneticness' in that some are more central to the system than the others and some are closer to the actual acoustico-articulatory medium? In that case, where do the fast-speech rules fit in, being only applied in certain situations, in fact sometimes undoing much of the work of our more abstract rules (wiping out long stressed vowels etc.). Do they perhaps form a special part of the phonology, 'the fast speech component'. I don't know.

FOOTNOTES

Chapter II

1. This rule is overridden in compound words, when the first syllable of the second part of the compound happens to be the second, fourth etc. syllable of the word as a whole: heimkynni, see Section 4.
2. A similar suggestion is made, according to Pétursson (1972), by A.S.Lieberman (Islandskaja prosodika...1971). Lieberman, it seems, connects length and stress in Icelandic, and considers the latter simply to be (in Pétursson's translation) the 'sommet quantitatif', the 'quantitative peak' of the syllable.
3. For the feature [+hard], see Chapter I, Section 3.
4. For an analysis along these lines see Árnason, forthcoming. For a completely different approach, where it is suggested that the preaspiration is a reflex of 'Accent 2', which shows up in most Norwegian and Swedish dialects, see Lieberman (1971). For suggestions of the analysis of preaspiration in terms of degree of glottal opening, see Þráinsson 1976.
5. I will not deal here with the proposals, made e.g. by Fudge (1969) and Sampson (1970) of 'generative phonotactics' with the syllable as an initial symbol of a rewrite grammar. This does not mean that I am sure that they are wrong, but there seem to me to be just as many problems facing proponents of that theory as those that believe in interpretive syllabification rules, and since there seem to be more people that believe in some kind of an interpretive syllabification theory, I will devote more attention to proposals within that theory.

6. There is, actually, a rather dubious assumption behind this, namely that [ø] is different from [œ] at the level on which syllabification takes place. There are good reasons to believe that [ø] and [œ] are allophones of the same phoneme (see Chapter I, Section 3). If we note this, we have to take account of the fact that [œj] and [œv] are permissible word-initial clusters: þjóð [œjou:ð] 'nation', þvo [œvɔ:] 'wash'. This may leave [ðj] and [ðv], phonologically /œj/ and /œv/, in exactly the same place as /sj/, /tj/ etc. and /nj/, /mj/ etc.
7. It is worth pointing out in connection with the mention of the law of finals and the law of initials, that the syllabification proposed by Garnes and Vennemann would break the law of initials, since forms like lausra [lœy:sra] 'loose' (adj. gen.pl.) would have to be syllabified lau-sra, giving syllable initial /sr/, which is not permitted word-initially, so their syllabification does not seem to be preferable on the grounds of the law of finals and law of initials.
8. I leave out of the discussion here what to do with intervocalic sequences of more than two consonants, mainly because they cause no problem as far as the length rule is concerned. At first glance it seems to me that they should be syllabified basically in the same way, that is with as many consonants as possible belonging to the stressed syllable, the exception being when the clusters end in /p, t, k, s/ + /v, j, r/. This would for example give ösk-ra, fölsk-vi, but ræksn-i, berkl-ar etc.

Chapter III

1. This 'overweight' on the second syllable is perhaps due to Accent 2, which in this area has a pitch peak on the second syllable as well as the first syllable (cf. Oftedal 1952:223).

Chapter IV

1. I am using the term 'prosodic' in a rather vague sense here, and perhaps wider than e.g. Lehiste (1970, cf. pp. 1-3) would subscribe to. I use it to cover things to do with length of syllables in Old Icelandic, which some might want to call phonotactic regularities. My use of the term is justified, I think, in the context of this study, since in Modern Icelandic at least the function of length is intimately connected with stress, as I hope is obvious from my discussion in Chapter II. It is only statements about syllable length in Old Icelandic which should perhaps be called phonotactic, but in connection with the changes that were to follow, I find it useful to apply the term prosodic to these phenomena as well as others that are perhaps more genuinely 'prosodic'.
2. It will have been noticed that the term syllable is here used in the maximal sense discussed above. This does not mean that the same syllabification principles were valid for Old Icelandic as Modern Icelandic (see Sections 2.3-4 of this chapter).
3. The Icelandic text is as follows (in a standardized orthography): Þat er leyfi háttanna, at hefa samstöfur seinar eða skjótar, svá at dragisk fram eða aptr ór rétttri tölum setningar, ok megu finnask svá seinar, at fimm samstöfur sé í þöru ok enu fyrir þá vísuorði, svá sem hér er:
4. The original text: Nú skal sýna svá skjótar samstöfur ok svá settar nær hverja annarri, at af því eykr lengó orðsins:
5. It is quite clear that it is an oversimplification to state an 'overnight hypothesis' and a 'conspiracy hypothesis' as the only two possible alternatives, since in actual fact linguistic changes need time to take place.

It seems that the lengthening of short vowels both in monosyllables and in bisyllabics took quite a long time to be completed, and it is very likely that the same was true of the shortening of long vowels. It may turn out that the difference between a single change or a conspiracy will be neutralized because of the 'gradualness' of the change(s). At this stage of the discussion, however, it seems profitable to make the simplifying assumption that the changes were or could have been more or less abrupt.

6. True, it is suggested by Sigmundsson (1970) that confusion in the spelling of intervocalic consonants (like berra for bera) that occurs in 14th and 15th century texts stems from changes that were taking place in the quantity structure. It is quite possible that spellings like these represent sporadic lengthening of consonants (cf. p. 305).
7. In this whole discussion on metrics, a few complications are left out concerning the metrical function of long and short syllables. For example, it seems that verbs could be metrically unstressed, and thus long verbal stems could function as short (metrically unstressed) as well as long. Another ambiguity in metrical function of linguistic forms is to be found in compounds, especially proper names like Ólafr (Áleifr), which could, it seems, either have a metrical function or , or even . Although these facts may make it necessary to make some qualifications to our statements concerning metrical rules, it does not affect our main concern here, that is, whether it seems that old long and short stressed syllables had different metrical functions.

8. The figures for Olgeirs rímur, Brávallarímur and Númarímur were obtained by a special survey. The material investigated was: Olgeirsrímur: Rímur no: I, II, IV, V, VII, VIII, IX, XI and XIII (Olgeirsrímur (1947) I:1-47), 1360 untruncated lines, 29 breaches; Brávallarímur: Rímur no. I, III IV, VIII. (Brávallarímur (1965):1-67), 572 untruncated lines, 47 breaches; Númarímur: The first ríma (Breiðfjörð 1937:7-20), 168 untruncated lines, 30 breaches.
9. It is perhaps worth considering whether these phenomena could be explained using Lass' (1971) idea of looking at boundaries as obstruents. One can for example say that word boundaries are like obstruents in that a consonant intervening between them and the preceding vowel will close a syllable. One can say that a word-final consonant will have to belong to the same syllable as the preceding vowel, since the word initial sequence $C\#$ is impermissible in most languages. There seems to me, however, to be something drastic about assigning the end-signals of words in some way a phonological status similar to regular segments. For one thing, boundaries have a very peculiar distribution. I would, then, like to find more positive evidence before using this interesting idea to explain metrical behaviour.
10. It seems to me that a 'non-realist' theory of language can only be saved from vacuity by claiming some metaphysical reality which acts as a referent for it. This could be something like an absolute notion of simplicity or harmony, which the different theories are evaluated against. It may be that this should be preferred to the realism described above. Yet it seems that a mentally real competence or a socially real language system, or even an abstract 'logic of language' are not as far out of reach as a universal rule of simplicity or 'world harmony'. Notice too that if some 'world harmony' is assumed, against which a theory of language, is ultimately to be tested, it is no longer 'non-realist' in a wider context

11. Jón Arason was the last Catholic bishop in Iceland, executed in 1550.
12. Perhaps one can say that structural changes are teleological, introduced in order to restore order to some more or less chaotic situation created by mechanical phonetic tendencies.

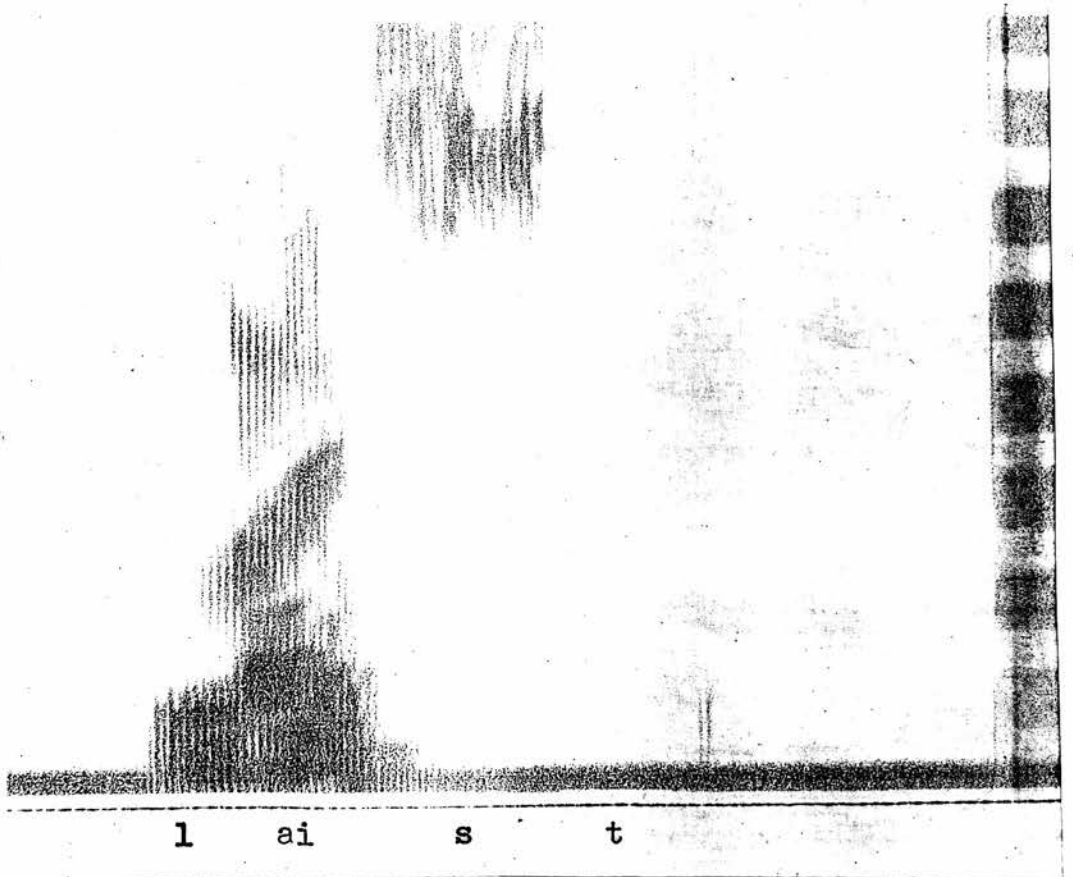
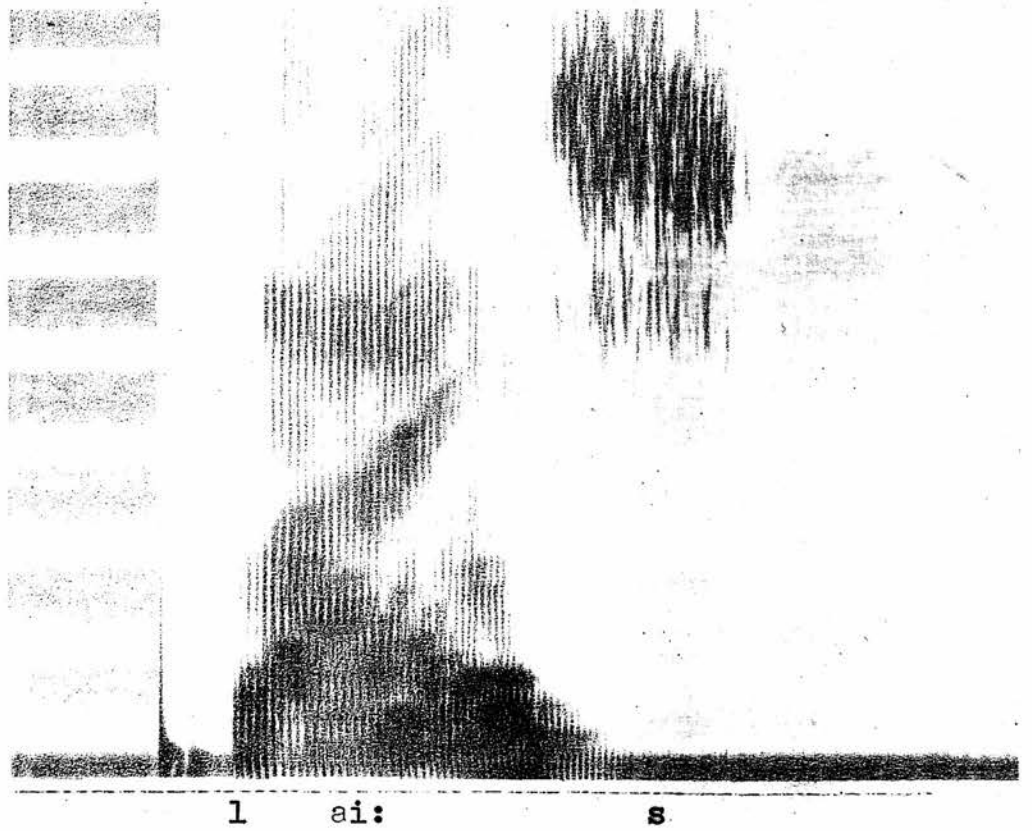
Chapter V

1. This whole discussion of distinctive features may seem a bit conceited and the assertions unsupported, but I wonder whether the general belief in universal phonological features is any better supported. It seems to me that it is largely based on the (severely challenged, cf. Derwing 1973 and Putnam 1971) assumption that children are equipped with a certain set of innate linguistic universals, among which are the phonological features, 30 or so (depending on your creed). It seems that these 'innate universals' would by nature have to be positive entities or concepts. But in actual languages the features are, I think, defined by their function within systems and largely negative. The only place where universality comes into the picture of distinctive features is in the actual physical scales on which the features operate. But these universal scales are non-linguistic, rather physical surroundings within which every human language works. Some physical phenomena may be better suited for use in language than others, and are therefore used more commonly than others, perhaps even in every language. (See Sampson (1974) for arguments against the existence of a 'universal phonetic alphabet', using data concerning tones.)
2. Phenomena of a similar sort in Lithuanian are described by Kenstowicz (1970). But Kenstowicz argues

that other phonological regularities can be accounted for more easily in terms of a feature analysis of length and suggests that both a 'sequence' analysis and a feature analysis are needed for Lithuanian.

3. It seems that Finnish is close to being a language of this kind. It has all vowels occurring as long or short, and there seems to be good reason to analyse long vowels as vowel geminates. The monophthongs are of eight qualities: a, e, i, o, u, y, ä, ö, and there are seventeen diphthongs, having each two consecutive qualities that can be identified with one of the simple vowels. (Fromm and Sadeniemi 1956:20-21). True, we are far from having the 56 theoretically possible combinations of the eight monophthongs, which indeed is not surprising (what would a language do with 64 vowel nuclei?), but the prediction of a relatively high number of diphthongs is at least in some way borne out by the facts. Some of the theoretically possible combinations of monophthongs into diphthongs are excluded automatically by a rule of vowel harmony, which demands that only front or only back vowels (with the exception of i and e, which stand outside the domain of the vowel-harmony rule) can cooccur within the same word. This excludes diphthongs like: aö, oä, uy, oy, ay, etc., altogether 18 diphthongs. Of the 38 remaining theoretically possible diphthongs, the following occur: ai, au, ei, eu, ey, ie, iu, oi, ou, ui, uo, yi, yö, äi, äy, öi, öy. It remains to be seen whether the absence of the other 21 theoretically possible diphthongs can be stated naturally in terms of phonotactic constraints, perhaps of the same type that make initial clusters like tk uncommon in languages.

4. It should perhaps be stressed, as must be evident from our discussion of abstractness, that these statements I have been making concerning the connection between geminateness of vowel length and 'pair-basedness' of relations between long and short vowel subsystems, should not be interpreted in the spirit of if-and-only-if statements. It may be that the grounds for a geminate-analysis are so strong (apart from the question of 'pair-basedness') that one would accept considerable ad hoc constraints against V - VV pairs for every vowel quality. I am only saying that, other things being equal, one should expect systems with length as geminateness to have a pair-based relation between two subsystems.
5. For a suggestion of a phonological analysis of the semivowels, see Benediktsson 1972:159-161 and references. Benediktsson suggests that syllabicity and length belong to the same underlying feature, that the rising diphthongs are clusters of a short vowel + a long vowel and the first vowel loses its syllabicity because of its shortness as opposed to the length of the second vowel. Thus ió would be underlying /io:/, and ei (falling diphthong) would be underlying /ei/. It seems to me that this analysis is problematic in that it does not account for the difference of já and ja ('long' and 'short' 'rising diphthong'.)
6. This is shown in spectrograms I have made of my own speech (cf. also Garnes 1974b). Although the quality movement in short diphthongs is often not very noticeable in listening, there is definitely a similarity in the movement of the vowel formants in pairs like læs: (læst, which must be assigned to the diphthongal nature of the ai - sound. The short variant looks like a miniature of the long one. (See spectrograms overleaf.)



Spectrograms, showing a long and a short [ai] in læs [lai:s] and last [laist] 'literate' (masc. vs. neut.).

BIBLIOGRAPHY

- Abercrombie, D. 1967. Elements of General Phonetics.
Edinburgh. Edinburgh University Press.
- Aitken, A.J. 1962. Vowel-Length in Modern Scots. Dupli-
cated. University of Edinburgh, Department of English^{Language}.
- Allen, W.S. 1965. Vox Latina: A Guide to the Pro-
nunciation of Classical Latin. Cambridge. Cambridge
University Press.
- Amundsen, S. 1964. Le vocalisme féroïen: Essai de phono-
logie diachronique. Fróðskaparrit. Annales Societatis
Scientiarum Færoensis 13:54-61.
- Andersen, H. 1972. Diphthongization. Language 48:11-50.
----- 1973. Abductive and deductive change.
Language 49:765-793.
- Andersen, P. 1954. Dansk fonetik. In N. Rh. Blegvad
(ed.): Nordisk lærebog for talepædagoger I. Copenhagen.
Rosenkilde og Baggers Forlag. (Pp. 308-353.)
- Anderson, J.M. 1969. Syllabic or non-syllabic phonology?
Journal of Linguistics 5:136-142.
----- 1975. Principles of Syllabification. York
Papers in Linguistics 5:7-20.
- Anderson, J.M and G. Jones 1974. Three theses concerning
phonological representations. Journal of Linguistics
10:1-26.
- Anderson, S.R. 1969. An outline of the phonology of
Modern Icelandic vowels. Foundations of Language 5:53-72.
----- 1972a. Icelandic y-umlaut and breaking in
a generative grammar. In E. Firchow, K. Grimsted, N.
Hasselmo, and W. O'Neil (eds.): Studies for Einar Haugen.
The Hague. Mouton.

- Anderson, S.R. 1972b. The Faroese Vowel System. In M. Brame (ed.): Contributions to Generative Phonology. Austin. University of Texas Press. (Pp. 1-21.)
- 1974. The Organization of Phonology. New York, San Francisco, London. Academic Press.
- 1976. Nasal consonants and the internal structure of segments. Language 52:326-344.
- Anttila, R. 1975. Exception as regularity in phonology. In W.U. Dressler and F.V. Mareš (eds.): Phonologica 1972. Akten der zweiten Internationalen Phonologie - Tagung. Wien, 5.-8. September 1972. München, Salzburg. Wilhelm Fink Verlag. (Pp. 91-99.)
- Árnason, K. 1974. Beygingarlýsing veikra segna. Mimeographed. Reykjavík. University of Iceland.
- 1975. Athugasemd um lengd hljóða í íslenzku. Mimir, blað stúdenta í íslenskum fræðum 23:15-19. (Reykjavík.)
- 1976. A note on Faroese vowels. Department of Linguistics, Edinburgh University: Work in Progress 9:58-63.
- Forthcoming (1977). Preaspiration in Modern Icelandic, phonetics and phonology. In a Festschrift for Jakob Benediktsson. Reykjavík.
- Bach, E. 1974. Explanatory inadequacy. In Cohen (ed.) 1974:153-171.
- Basbøll, H. 1970-71. A Commentary on Hjelmslev's Outline of the Danish Expression System (I). Acta Linguistica Hafniensia XIII.:173-211.
- 1974. The phonological syllable with special reference to Danish. Annual Report of the Institute of Phonetics, University of Copenhagen 8:39-128.

- Benediktsson, H. 1959. The Vowel System of Icelandic: A Survey of its History. Word 15:282-312.
- 1962. The Unstressed and the Non-Syllabic Vowels of Old Icelandic. Arkiv för nordisk filologi LXXVII:7-31.
- 1963. The Non-Uniqueness of Phonemic Solutions: Quantity and Stress in Icelandic. Phonetica 10:133-153.
- 1964. Old Norse Short g: One Phoneme or Two? Arkiv för nordisk filologi LXXIX:63-104.
- 1965. Early Icelandic Script as Illustrated in Vernacular Texts from the Twelfth and Thirteenth Centuries. Reykjavik. The Manuscript Institute of Iceland.
- 1968. Indirect changes of phonological structure: Nordic vowel quantity. Acta Linguistica Hafniensia XI:31-65.
- 1972 (ed.). The First Grammatical Treatise. University of Iceland. Publications in Linguistics I. Reykjavik. Institute of Nordic Linguistics.
- Bergman, G. 1973. A Short History of the Swedish Language. Translated by Francis P. Magoun jr. and Helge Kökeritz. Lund. The Swedish Institute for Cultural Relations with Foreign Countries. (Second revised edition.)
- Bergsveinsson, S. 1941. Grundfragen der isländischen Satzphonetik. Berlin, Copenhagen. Metten & Co. - Ejnar Munksgaard.
- Bjerrum, M. 1964. Forsøg til en analyse af det færøske udtrykssystem. Acta Philologica Scandinavica 25:31-69.
- Bloomfield, L. 1933. Language. New York, Chicago, San Francisco, Toronto. Holt, Rinehart and Winston.
- Botha, R.P. 1971. Methodological Aspects of Transformational Generative Phonology. Janua Linguarum. Series Minor 112. The Hague, Paris. Mouton.

- Brávallarímur 1965. Brávallarímur eftir Árna Böðvarsson.
(Edited by Björn K. Þórólfsson.) Rit Rímnafélagsins
VIII. Reykjavík. Rímnafélagið.
- Breiðfjörð, S. 1937. Núma rímur (3rd edition.) Reykjavík.
Snebjörn Jónsson, The English Bookshop. (First published
1833.)
- Cathey, J.E. and R.A. Demers 1976. On establishing
linguistic universals: A case for in-depth synchronic
analysis. Language 52:611-630.
- Chapman, K.G. 1962. Icelandic - Norwegian Linguistic
Relationships. Norsk tidsskrift for sprogvidenskap. Suppl.
bind 7. Oslo.
- Chen, M. 1976. Relative chronology: three methods of re-
construction. Journal of Linguistics 12:209-258.
- Chomsky, N. 1965. Aspects of the Theory of Syntax.
Cambridge, Mass. M.I.T. Press.
- Chomsky, N. and M. Halle 1968. The Sound Pattern of
English. New York, Evanston and London. Harper & Row,
Publishers.
- Christiansen, H. 1946-48. Norske dialekter. Oslo.
Johan Grundt Tanum.
- Cohen, D. 1974 (ed.). Explaining Linguistic Phenomena.
Washington D.C. Hemisphere Publishing Corporation.
- Craigie, W.A. 1900. On some Points in Skaldic Metre.
Arkiv för nordisk filologi, XVI:341-384.
- 1952. Sýnisbók íslenskra rímna I-III.
(Specimens of Icelandic Rímur. With introductions on the
history, metres and language of rímur.) London, Edinburgh,
Paris, Melbourne, Toronto and New York. Thomas Nelson
and Sons, Ltd.
- Derwing, B.L. 1973. Transformational grammar as a theory
of language acquisition. Cambridge. Cambridge University
Press.

- Dobson, E.J. 1962. Middle English Lengthening in Open Syllables. Transactions of the Philological Society 1962:124-148.
- Dressler, W.U. 1975. Methodisches zu Allegro-Regeln. In W.U. Dressler and F.V. Mareš (eds.): Phonologica 1972. Akten der zweiten Internationalen Phonologie - Tagung. Wien. 5.-8. September 1972. München, Salzburg. Wilhelm Fink Verlag. (Pp. 219-234)
- Durand, M. 1939. Durée phonétique et durée phonologique. In E. Blancquaert and W. Pée (eds.): Proceedings of the Third International Congress of Phonetic Sciences. Ghent. Laboratory of Phonetics of the University. (Pp. 261-265.)
- Einarsson, S. 1927. Beiträge zur Phonetik der isländischen Sprache. Oslo. A.W. Brøggers Boktrykkeri A/S.
- 1945. Icelandic: Grammar, texts, glossary. Baltimore. Johns Hopkins Press.
- Elert, C.-C. 1964. Phonologic Studies of Quantity in Swedish. Based on Material from Stockholm Speakers. Monografier utgivna af Stockholms Kommunalförvaltning. Uppsala.
- Eliasson, S. and N. La Pelle 1973. Generativa regler för svenskans kvantitet. Arkiv för nordisk filologi LXXXVIII: 133-148.
- Fromm, H. and M. Sadeniemi 1956. Finnisches Elementarbuch I. Grammatik. Heidelberg. Carl Winter - Universitätsverlag.
- Fudge, E.C. 1969. Syllables. Journal of Linguistics 5: 253-286.
- Garnes, S. 1973. Phonetic evidence supporting a phonological analysis. Journal of Phonetics 1:273-283.
- 1974a. Quantity in Icelandic: Production and Perception. Ph. D. dissertation. Ohio State University.
- 1974b. Suprasegmental Aspects of Icelandic Vowel Quality. Mimeographed. Ohio State University.

Garnes, S. 1975a. Perception, Production and Language Change. To appear in Papers from the Parasession on Functionalism, Chicago Linguistic Society. (Mimeographed.)

----- 1975b. The Nordic Quantity Shift: Quantity vs. Quality. Paper presented at the Winter Meeting of the Linguistic Society of America. San Francisco, December 1975. (Mimeographed.)

Gårding, E. 1973. The Scandinavian word accents. Phonetics Laboratory, Lund University. Working Papers 8.

Grundt, A. 1973. Open Syllable Lengthening in English: A Study in Compensatory Phonological Processes. Ph. D. dissertation. University of California, Berkeley.

Guðfinnsson, B. 1946. Mállýskur I. Reykjavík. Ísefoldarprentsmiðja.

----- 1964. Um íslenzken fremburð. Mállýskur II. (Ed. by Ólafur M. Ólafsson and Óskar Ó. Halldórsson.) Studia islandica 23. Reykjavík. Heimspekideild Háskóla Íslands - Bókaútgáfa Menningarsjóðs.

Halle, M. 1959/71. The Sound Pattern of Russian: A Linguistic and Acoustical Investigation. Second printing. The Hague, Paris. Mouton.

Hemmersheimb, V.U. 1891. Færósk antologi I: Tekst samt historisk og grammatisk indledning. Copenhagen. S.L. Møllers Bogtrykkeri (Møller & Thomsen).

Harris, J.W. 1969. Sound change in Spanish and the theory of markedness. Language 45:538-552.

Haugen, E. 1949. Phoneme or prosodeme. Language 25:278-282.

----- 1958. The phonemics of Modern Icelandic. Language 34:55-88.

----- 1967. On the rules of Norwegian tonality. Language 43:185-202.

- Hempel, C.G. 1966. Philosophy of Natural Science. Englewood Cliffs, N.J. Prentice Hall Inc.
- Hjelmslev, L. 1951/1973. Outline of the Danish Expression System with Special Reference to the Stød. In Essais linguistiques par Louis Hjelmslev II. Travaux du Cercle linguistique de Copenhague. Vol. XIV. (1973). Copenhagen. Nordisk Sprog-og Kulturforlag. (Pp. 247-266.)
- Hoard, J.E. 1971. Aspiration, tenseness, and syllabication in English. Language 47:133-140.
- Hooper, J.B. 1972. The syllable in phonological theory. Language 48:525-540.
- 1976. Constraints on Schwa-deletion in English. Paper presented at the International Conference on Historical Phonology, Ustronie, Poland. March 17.-20., 1976.
- Indrebø, G. 1951. Norsk målsoga. Ed. by Per Hovda and Per Thorson. Bergen. A.S. John Griegs Boktrykkeri.
- Jakobson, R., G. Fant and M. Halle 1952. Preliminaries to Speech Analysis. M.I.T. Acoustics Laboratories, Technical Report No. 13. Cambridge, Massachusetts. The M.I.T. Press.
- Jakobson, R. and M. Halle 1956. Fundamentals of Language. Janua Linguarum. Series Minor 1. The Hague. Mouton.
- Jeffers, R. J. 1974. On the notion 'explanation' in historical linguistics. In J.M. Anderson and C. Jones (eds.): Historical Linguistics II. Theory and description in phonology. North-Holland Linguistic Series 12b. Amsterdam, Oxford. North-Holland Publishing Company. (Pp. 231-255.)
- Jespersen, O. 1909/1961. A Modern English Grammar on Historical Principles. Part I: Sounds and Spellings. London. George Allen and Unwin Ltd.; Copenhagen. Ejnar Munksgaard.
- 1922. Modersmålets fonetik. Copenhagen and Oslo (Kristiania). Gyldendalske Boghandel - Nordisk Forlag.

- Jones, D. 1962. The Phoneme: its Nature and Use. Cambridge. W. Heffer and Sons Ltd.
- Jónsson, F. 1912-15. Den norsk-islandske skjaldedigtning, udgivet af Kommissionen for det Arnamagnæanske legat ved Finnur Jónsson. AI-II, BI-II. Copenhagen and Kristiania (Oslo). Gyldendalske Boghandel - Nordisk Forlag.
- Kahn, D. 1976. Syllable-based Generalizations in English Phonology. M.I.T. dissertation. Reproduced by the Indiana University Linguistics Club. Bloomington, Indiana.
- Karlsson, S. 1964. Gömul hljóðvöl í ungum rímum. Lingua Islandica - Íslensk tunga V:7-29.
- Keller, W. 1908. Die Akzente in den ags. Handschriften. Prager deutsche Studien VIII. (Pp.97-120.)
- Keller, R.E. 1961. German dialects: phonology and morphology. Manchester. Manchester University Press.
- Kenstowicz, M. 1970. On the Notation of Vowel Length in Lithuanian. Papers in Linguistics 3.1:73-113.
- Kelly, J. and H. Kress 1972. Icelandic: Readings, Grammar, Exercises. Mimeographed. Reykjavik. University of Iceland.
- Von Kienle, R. 1960. Historische Laut- und Formenlehre des Deutschen. Tübingen. Max Niemeyer Verlag.
- King, R.D. 1974. Can Rules be Added in the Middle of Grammars? Reproduced by the Indiana University Linguistics Club. Bloomington, Indiana.
- Kiparsky, P. 1968a. Linguistic universals and linguistic change. In E. Bach and R.T. Harms (eds.): Universals in linguistic theory. London, New York, Sydney, Toronto. Holt, Rinehart and Winston. (Pp. 170-202.)
- 1968b. How Abstract is Phonology? Reproduced by the Indiana University Linguistics Club. Bloomington, Indiana.

- Kisseberth, G. W. 1970. On the Functional Unity of Phonological Rules. Linguistic Inquiry 1:291-306.
- Kjartansson, H.S. 1971. Hjóðdvalarbreytingin, lítil athugasemd um heimildargildi kveðskapar. Mimir, blað stúdenta í íslenskum fræðum 10. (Reykjavík.)
- 1974. Lengd íslenzkra sérhljóða. Mimir, blað stúdenta í íslenskum fræðum 22. (Reykjavík.)
- Króka-Refs saga og Króka-Refs rímur. Udgivne af Pálmi Pálsson. Samfund til udgivelse af gammel nordisk litteratur X. Copenhagen. S.L. Møllers Bogtrykkeri. (1883.)
- Labov, W. 1965. On the mechanism of linguistic change. In C.W. Kreidler (ed.): Report of the Sixteenth Annual Round Table Meeting on Linguistics and Language Studies, Georgetown University Monograph Series on Languages and Linguistics 18. Washington, D.C. Georgetown University Press. (Pp. 91-114.)
- Ladefoged, P. 1971. Preliminaries to Linguistic Phonetics. Chicago, London. The University of Chicago Press.
- Langendoen, D.T. 1976. Review of Cohen (1974). Language 52:690-695.
- Lass, R. 1971. Boundaries as obstruents: Old English voicing assimilation and universal strength hierarchies. Journal of Linguistics 7:15-30.
- 1974. Linguistic orthogenesis? Scots vowel quantity and the English length conspiracy. In J.M. Anderson and C. Jones (eds.): Historical linguistics II. Theory and description in phonology. North-Holland Linguistic Series 12b. Amsterdam, Oxford. North-Holland Publishing Company. (Pp. 311-352.)
- 1976a. English phonology and phonological theory: synchronic and diachronic studies. Cambridge. Cambridge University Press.
- 1976b. On generative taxonomy, and whether formalizations 'explain'. Studia Linguistica XXX:139-154.

- Lass, R. and J.M. Anderson 1975. Old English phonology. Cambridge. Cambridge University Press.
- Lehiste, I. 1970. Suprasegmentals. Cambridge, Mass., London. The M.I.T. Press.
- Liberman, A.S. 1971. Gibt er Silbenakzente im Isländischen? Nordeuropa. Studien 4:173-183. (Greifswald.)
- Linell, P. 1974. Problems of psychological reality in generative phonology. Reports from Uppsala University Department of Linguistics 4. Uppsala.
- Lockwood, W.B. 1955. An Introduction to Modern Faroese. Copenhagen. Ejnar Munksgaard.
- Melone, K. 1952. The Phonemes of Modern Icelandic. In: Studies in Honor of Albert Morey Sturtevant. Lawrence. University of Kansas Press. (Pp. 5-21.)
- 1953. Long and short in Icelandic phonemics. Language 29:61-62.
- Marchand, J.W. 1973. The Sounds and Phonemes of Wulfila's Gothic. Janua Linguarum. Series Practica 25. The Hague, Paris. Mouton.
- Von Mises, R. 1951. Positivism: a study in human understanding. New York. Dover Publications.
- Mitzka, W. 1954. Die dänische und die deutsche Konsonantenschwächung. Zeitschrift für Mundartforschung XXII:65-87.
- Noreen, A. 1904. Altnordische Grammatik II. Altschwedische Grammatik mit Einschluss des Altgutnischen. Halle. Max Niemeyer.
- 1923/1970. Altnordische Grammatik I. Altisländische und altnorwegische Grammatik. 5., unveränderte Auflage. Tübingen. Max Niemeyer Verlag. (4th Edition 1923.)

Ófeigsson, J. 1920-24. Træk af moderne islandsk Lydlære. In S. Blöndal: Islandsk-dansk Ordbog. (Íslenszk-dönsk orðabók). Reykjavík. Verslun Þórarins B. Þorlákssonar.- Copenhagen. H. Aschehoug & Co. (W. Nygaard). (Pp. XIV-XXVII.)

Oftedal, M. 1952. On the Origin of the Scandinavian Tone Distinction. Norsk tidsskrift for sprogvidenskab XVI:201-225.

Olgeirs rímur 1947. Olgeirs rímur danska eftir Guðmund Bergþórsson. I-II. (Edited by Björn K. Þórólfsson and Finnur Sigmundsson.) Reykjavík. Landsbókasafn Íslands - Ísafoldarprentsmiðja.

O'Neil, W.A. 1964. Faroese vowel morphophonemics. Language 40:366-371.

Orešnik, J. 1971. On the phonological boundary between constituents of Modern Icelandic compound words. Linguistica XI:51-59. (Ljubljana.)

----- 1972. On the Epenthesis Rule in Modern Icelandic. Arkiv för nordisk filologi LXXXVII:1-32.

Paul, H./W. Mitzka 1963. Mittelhochdeutsche Grammatik. 19. Auflage. Tübingen. Max Niemeyer Verlag.

Pétursson, M. 1972. Review of: A.S. Liberman: Íslandskaja prosodika. K fonologičeskoj xarakteristike sovremennogo islandskogo jazyka i ego istorii. (Leningrad 1971.) Phonetica 26:89-128.

----- 1974. Les articulations de l'islandais a la lumière de la radiocinématographie. Paris. Librairie C. Klincksieck.

Pontus rímur 1961. Pontus rímur eftir Magnús Jónsson, þrúða, Pétur Einarsson og síra Ólaf Halldórsson. (Edited by Grímur M. Helgason.) Rit Rímnafélagsins X. Reykjavík Rímnafélagið.

Popper, K.R. 1968. Conjectures and Refutations. The Growth of Scientific Knowledge. London and Henley. Routledge and Kegan Paul.

Prokosch, E. 1939. A Comparative Germanic Grammar. Baltimore. Linguistic Society of America.

Pulgram, E. 1970. Syllable, Word, Nexus, Cursus. Janua Linguarum. Series Minor 81. The Hague, Paris. Mouton.

Putnam, H. ¹⁹⁷¹ The 'Innateness Hypothesis' and Explanatory Models in Linguistics. In J.R. Searle (ed.): The Philosophy of Language. Oxford. The Oxford University Press. (Pp. 130-139.)

Ralph, B. 1975. Phonological Differentiation: Studies in Nordic Language History. Nordistica Gothoburgensia 8. Göteborg. Acta Universitatis Gothoburgensis.

Rasmussen, P. 1972. Översigt over det danske udtrykssystemets historie. Mimeographed. Reykjavík. Bóksala stúdenta.

Rímnaefn I-II. Samling af de ældste islandske rimer. Udgivet for Samfund til udgivelse af gammel nordisk litteratur (XXXV) ved Finnur Jónsson. Copenhagen. S.L. Møllers Bogtrykkeri. (1905-1922.)

Rischel, J. 1964. Toward the phonetic description of Faroese vowels. Fróðskaparrit. Annales Societatis Scientiarum Færoensis 13:99-113.

----- 1968. Diphthongization in Faroese. Acta Linguistica Hafniensia XI:89-118.

Sampson, G. 1970. On the need for a phonological base. Language 46:586-626.

----- 1974. Is there a universal phonetic alphabet? Language 50:236-259.

Scriven, M. 1959. Explanation and Prediction in Evolutionary Theory. Science 130:477-482.

Sievers, E. 1893. Altgermanische Metrik. Halle. Max Niemeyer Verlag.

- Sigmundsson, S. 1970. Um hljóðvöl í íslenzku. Fróðskaparrit. Annales Societatis Scientiarum Færoensis 18: 320-332.
- Sigurd, B. 1955. Rank order of consonants established by distributional criteria. Studia Linguistica IX:8-20.
- Skautrup, P. 1944. Det danske sprogs historie. Første bind. Fra Guldhornene til Skånske lov. Copenhagen. Gyldendalske Boghandel-Nordisk Forlag.
- Snorra Edda 1931. Edda Snorra Sturlusonar udgivet efter håndskrifterne af Kommissionen for det Arnamagnæanske legat ved Finnur Jónsson. Copenhagen. Gyldendalske Boghandel - Nordisk Forlag.
- Söderström, S. 1972. Om kvantitetsutvecklingen i norrländska folkmål. Acta Academiae Regiae Gustavi Adolphi LII. Uppsala. Lundequistska Bokhandeln.
- Sommerfelt, A. 1951/1962. The development of quantity as evidence of Western European linguistic interdependence. In A. Sommerfelt: Diachronic and Synchronic Aspects of Language. Selected articles. Janua Linguarum. Series Maior VII.'S -Gravenhage. Mouton. (Pp. 81-86. First published in: English Studies To-Day, 1951.)
- Steblin-Kamenskij, M.I. 1960. The vowel system of Modern Icelandic. Studia Linguistica XIV:35-46.
- Taylor, J.E. 1973. A Generative Phonology of Faroese, Utilizing Unordered Rules. Ph. D. Dissertation. Indiana University.
- Trubetzkoy, N.S. 1958. Grundzüge der Phonologie. 2. Auflage. Göttingen. Vandenhoeck & Ruprecht.
- Vacheck, J. 1959. Notes on the Quantitative Correlation of Vowels in the Phonematic Development of English. Mélanges de linguistique et de philologie. Fernand Mossé in memoriam. Paris. Didier. (Pp. 444-456.)
- Vago, R.M. 1977. In Support of extrinsic ordering. Journal of Linguistics 13:25-41.

Vennemann, T. 1971. The phonology of Gothic vowels. Language 47:90-132.

----- 1972. On the Theory of Syllabic Phonology. Linguistische Berichte 18:1-18.

Wang, W.S -Y. 1969. Competing changes as a cause of residue. Language 45:9-25.

Weinreich, U., W. Labov and M.I. Herzog 1968. Empirical Foundations for a Theory of Language Change. In W.P. Lehmann and Y. Malkiel (eds.): Directions for Historical Linguistics. A Symposium. Austin, London. University of Texas Press. (Pp. 95-195.)

Weinstock, J. 1975. Quantity and Labialization in the Nordic Languages: An Historical Interpretation. In K.-H. Dahlstedt (ed.): The Nordic Languages and Modern Linguistics 2. Stockholm. Almqvist & Wiksell International. (Pp. 756-773).

Wessén, E. 1945. Svensk språkhistoria I. Ljudlära och ordböjningslära. (2nd edition). Stockholm. Filologiska föreningen vid Stockholms Högskola.

----- 1960. Våra folkmål. (6th edition). Malmö. Fritzes. (1st edition 1935.)

Woo, N. 1972. Prosody and Phonology. Reproduced by the Indiana University Linguistics Club. Bloomington, Indiana.

Zachariassen, U. 1968. Munurin millum Suðuroyarmál og fýroyskt annars í longdaruppfatanini av p, t, k, s, + j, r og p, k + l. Fróðskaparrit. Annales Societatis Scientiarum Færoensis 16:45-51.

Zwicky, A.M. 1972. Note on a phonological hierarchy in English. In R.P. Stockwell and R.K.S. Macaulay (eds.): Linguistic Change and Generative Theory. Bloomington and London. Indiana University Press. (Pp. 275-301.)

Pórólfsson, B.K. 1925. Um íslenskar orömyndir á 14. og 15. öld og breytingar þeirra úr fornmálinu. Reykjavík. Fjelagsprentsmiðjan.

----- 1929a. Kvantitetsomvæltningin í íslensk. Arkiv för nordisk filologi XLV:35-81.

----- 1929b. Nokkur orð um hinar íslensku hljóðbreytingar á > ie og ɣ, ʏ, ey > i, í, ei. In Studier tillägnade Axel Kock. Supplement to Arkiv för nordisk filologi XL:232-243. Lund. C.W.K. Gleerup.

----- 1934. Rimur fyrir 1600. Safn Fræðafjelagsins um Ísland og Íslendinga IX. Copenhagen. S.L. Möller.

Þorsteinsson, B. 1906-1909. Íslensk bjóðlög. Copenhagen. S.L. Möller.

Þráinsson (Þhráinsson) H., 1976. Dialectal variation in Icelandic as evidence for aspiration theories. Paper presented at the Third International Conference on Nordic and General Linguistics. Austin, Texas. April 1976.