STUDIA PHILOLOGICA JYVÄSKYLÄENSIA VI

ASPECTS OF QUANTITY IN STANDARD FINNISH

JAAKKO LEHTONEN

JYVÄSKYLÄ 1970

STUDIA PHILOLOGICA JYVÄSKYLÄENSIA VI

ASPECTS OF QUANTITY IN STANDARD FINNISH

JAAKKO LEHTONEN

ACADEMIC DISSERTATION TO BE PUBLICLY DISCUSSED, BY PERMISSION OF THE FACULTY OF HUMANITIES, UNIVERSITY OF JYVÄSKYLÄ, IN AUDITORIUM II–212, ON SEPTEMBER 26, 1970, AT 12 O'CLOCK NOON

JYVÄSKYLÄ 1970

ASPECTS OF QUANTITY IN STANDARD FINNISH

STUDIA PHILOLOGICA JYVÄSKYLÄENSIA VI

ASPECTS OF QUANTITY IN STANDARD FINNISH

JAAKKO LEHTONEN

JYVÄSKYLÄ 1970

KUSTANTAJA
JYVÄSKYLÄN YLIOPISTO

URN:ISBN:978-951-39-8605-6
ISBN 978-951-39-8605-6 (PDF)
ISSN 0585-5462

K. J. Gummerus Osakeyhtiön kirjapainossa
Jyväskylässä 1970

CONTENTS

page

PREFACE 9

1. INTRODUCTION 11

1. 1. *The concepts of quantity, length and duration* 13

1. 2. *Interference in physical characteristics, perceptual parametres and linguistic functions* 17

2. THE FUNCTION OF QUANTITY IN FINNISH 24

2. 1. *The phonemic structure of Finnish* 24

2. 1. 1. Segmental phonemes 24

2. 1. 2. Some phonotactic restrictions 26

2. 1. 3. Phonemic word structures 28

2. 1. 4. The number of word structures 29

2. 2. *Phonemic interpretation of quantities* 30

3. THE ANALYSIS OF PHONETIC DURATIONS 37

3. 1. *Concept of standard language and the geographical position of Jyväskylä town* 37

3. 2. *Material for the analysis* 38

3. 2. 1. Selection of informants 38

3. 2. 2. Reading of test sentences 41

3. 2. 3. Durational differences between speakers 41

3. 2. 4. Linguistic material 43

3. 3. *Technical procedures* 45

3. 3. 1. »The segmentator» 47

3. 4. *Definition of phonetic segments* 49

3. 4. 1. Description of segment boundaries 50

3. 4. 1. 1. Plosives 51

3. 4. 1. 2. Fricatives /s/ and /h/ 53

3. 4. 1. 3. Nasals /m/, /n/ and /ŋ/ 54

3. 4. 1. 4. Semi-vowels /v/ and /j/ 56
3. 4. 1. 5. Liquids /l/ and /r/ 57
3. 5. *Statistical description and analysis* 58
3. 6. *Intrinsic duration characteristics of segmental phoneme units* 60
3. 6. 1. Vowel durations 62
3. 6. 1. 1. Duration of dihthongs 69
3. 6. 2. Duration of consonants 70
3. 6. 3. The effect of adjacent sounds on sound duration 76
3. 6. 3. 1. The effect of preceding and following consonants on the duration of a vowel 76
3. 6. 3. 1. 1. The effect of word-initial consonant on vowel duration 76
3. 6. 3. 1. 2. The effect on vowel duration of a consonant following a vowel 77
3. 6. 3. 1. 3. The effect of the final consonant of a word on the durations of the preceding sound segments 81
3. 6. 3. 2. The effect of vowels on the duration of consonants 85
3. 7. *Durations of sound segments in various phoneme syntagms* 86
3. 7. 1. Durational distinction between short and long vowel quantity 86
3. 7. 2. Durational C/CC distinction of consonants 91
3. 7. 2. 1. Quantity and the Anschluss concept 91
3. 7. 2. 2. The concept of geminate consonants 93
3. 7. 2. 3. The durational distinction between different single and geminate consonants 96
3. 7. 2. 4. Consonant clusters 99
3. 7. 3. Quantity distinction and the commutation procedure 102
3. 7. 4. Comparisons (1.–12.) of segmental durations in minimally distinctive word structure pairs 105
3. 7. 4. 1. Synopsis of comparisons 122
3. 7. 4. 2. The influence on the environment of the length of the first syllable vowel 122
3. 7. 4. 3. The influence of the length of a second syllable vowel on the environment 123

3.7.4.4. The influence of an intervocal consonant on the environment 124
3.7.5. A general review of bisyllabic structures 125
3.7.6. General review of three-syllable structures 120
3.8. *The half-long vowel* 132
3.8.1. The effect of the half-long vowel on the phonetic distinctions of structures 134
3.8.2. Diachronic survey 136
3.9. *The effect of word length on sound duration* 138
3.10. *The effect of stress on the duration of sounds* 145
3.11. *Durational distribution in slow speech* 147
3.12. *Units of temporal organization of sound sequences* 148
3.12.1. Speech measure and measure 149
3.12.2. Syllable and mora 150

4. PERCEPTION TESTS 153
4.1. *Arrangement of the listening tests* 155
4.2. *Shortening of word-initial vowels* 157
4.3. *Shortening of second syllable vowels* 164
4.3.1. Shortening of the second syllable vowel of the original CVCVV words 164
4.3.2. Shortening of the original second syllable double vowel after a geminate consonant 167
4.4. *The changes in the duration of intervocal consonants* 171
4.4.1. The changes in the duration of the intervocal single consonant 172
4.4.2. Durational changes in the intervocal consonant of the original CVCCV words 177
4.4.3. Shortening of an original double consonant preceding a second syllable double vowel 183
4.5. *Conclusions and assessments* 185
4.6. *The bandwidth of phonemic boundary* 187
4.7. *Theory of the perception of phonemic length* 189

SUMMARY AND OUTLOOKS 192

REFERENCES 195

PREFACE

My interest in phonetics and phonemics, the experimental treatment and functional appraisal of the various phenomena of spoken language, derives from the stimulating lectures and work of my teacher of Finnish, Professor Aarni Penttilä, who was also responsible for setting up the phonetics laboratory at the University of Jyväskylä where the instrumental analysis for the present work was performed.

The inspiration for this study is to be found in the fundamental works of Professor Kalevi Wiik in the field of phonetics and phonemics; in his work he has also dealt with problems of quantity. Professor Wiik's favourable attitude to my research schemes has been the most valuable support I have had during my work. To him I wish to express my sincere thanks for his advice and suggestions concerning the problems met with in my study.

My thanks are also due to Professor Heikki Leskinen, Head of the Department of Finnish at the University of Jyväskylä, for his encouragement, advice and comments during various phases of my work.

I should also like to take this opportunity to express my gratitude to the following persons and departments, without whose help this study would never have been possible: all my informants; the principals of the Jyväskylän Lyseo and Jyväskylän Tyttölyseo for their help in the arrangement of the listening tests; the Computer Centre of the University of Jyväskylä where the results were analysed; Mr Keijo Lipén who supervised the calculations and patiently instructed me in the statistical analysis of the material; Mr Aaro Sorsa, for his special experience in the solution of the many theoretical and practical problems concerning the construction of the analysis equipment, especially the segmentator; my colleagues, and in particular Mr Leho Võrk and Mr Fred Karlsson, for the many stimulating discussions on the topics of my investigation which were of inestimable value in the formation and clarification of my ideas; and my family and relatives for their patience and support during my work.

The late Mr John Stratton undertook the arduous task of translating the manuscript. I must record here the debt of gratitude I owe him and which unhappily I can no longer express to him.

For the financial support I have received I am indebted to the University of Jyväskylä and the Suomen Kulttuurirahasto Fund, which enabled me to complete my work.

Finally I should like to express my debt to the University of Jyväskylä for accepting this work as part of its series, Studia Philologica Jyväskyläensia.

Jyväskylä, September 1970
Jaakko Lehtonen

1. INTRODUCTION

The purpose of the present work is to illustrate the durational patterns of Standard Finnish in the light of the kind of educated Finnish spoken by native Jyväskylä residents. The town Jyväskylä is not in any special position in respect of the standard form of the Finnish language. It is not presumed that the educated spoken Finnish here would be more pure or less influenced by the regional dialect background than somewhere else. It is likely that the surrounding dialect everywhere affects the educated standard speech too. However, in the opinion of the present author, the timing differences which are found in the educated speech of people of different dialect-geographical background do not affect the general character of the Finnish quantity patterns described in this study.

The analysis is based on segmentation of the acoustic output of speech and the results of analysis are supplemented and amended by identification tests performed with artificially modified speech stimuli. In the first chapters the ambiguity of the traditional quantity concept and the subjective evaluation of phonetic duration are discussed. A normal, phonetically untrained Finnish listener can only perceive one 'short' quantity in normal spoken Finnish, even though a 'short' vowel in one position may be more than two times as long in duration as in another position. The traditional quantity concept is clarified by dividing it into three: duration, length and subjective quantity. While among the latter concepts length is identified as a distinctive category, subjective quantity denotes merely the subjective impression of articulation time. The interference of different physical and perceptual parametres and the support of various parametres on quantity perception are also discussed in the first chapters.

In the phonemization of the Finnish language made in accordance with so-called autonomous or taxonomic phonemics the long quantities of both vowels and consonants are interpreted as combinations of two identical vowel or consonant phonemes and equalized with combinations of two dissimilar vowel phonemes (diphthongs) and clusters of dissimilar consonants. As one possibility for the description of prosodic length patterns the mora concept is introduced.

In the first part of the experimental work proper the various factors affecting the duration of a given phonetic sound segment are discussed. These factors can roughly be divided into two: factors related to the special properties of each segmental phoneme and factors related to the position of the sound segment in a larger phonemic entity. The former are investigated by comparing the absolute and intrinsic durations of different vowels and consonants, mainly by means of various nonsense-words. Besides the durational values of the different speech-sounds themselves, the effects of the different sounds on the sound segments surrounding them are also discussed. It is possible that a great part of the durational differences between various vowels and respectively between consonants, and of the differences in their effect on neighbouring segments are due to some common motor and physiological properties of the articulation mechanism. However, there is no reason to claim that the same factors would not have any effect on the syntagmatic durational patterns at all.

The distribution of duration in various phoneme syntagms is tested by comparing the segmental durations in minimally differing word structure series. In most of the compared pairs the difference between structures is situated in a larger domain than the segment which is described as distinctive on the phonemic level. The most noticeable differences of this kind are caused by the existence of the so-called half-long vowel, which is also discussed separately.

The purpose of the second part of the work is to study the function of the durational patterns revealed by the analysis from the point of view of perception. The identification of the structure of a word (or perhaps of a measure) seems to be determined on the basis of two kinds of contrasts: a contrast between the durations of vowels in successive syllables and a contrast of a consonant with both the preceding and the following vowel segment. On the level of acoustic output of speech there are difficulties, especially in the defining of the latter ratio, because of the differences in the inherent duration values of different consonants and in their effects on the duration of surrounding vowels. One has to assume that a listener is able to amend the durational pattern of the output by the 'coefficients' of the various factors influencing the distribution of segmental duration and that he is thus able to uncover the underlying timing pattern of the intended structure.

It seems possible that in the 'deep structure' of speech production, perhaps on some early neurological level, the temporal structure of a phonemic word corresponds to the phonemic composition of the word in that a long segment equals the sequence of two identical short segments and that all of the short or single segments are of equal value regardless of their position. These sequences are however modified by the timing factors of measure (perhaps of word) and of stress, intonation and those of the properties of different qualitative units.

1. 1. *The concepts of quantity, length and duration*

In everyday non-phonetic linguistic usage the concept of quantity does not seem to present any problems: a segmental phone or other unit being described which, according to the interpretation of the observer, lasts longer or takes more time, is described as having a longer quantity than one which has, according to observations, required less time in pronunciation. The researcher can divide the quantities he has to estimate into two or more categories called degrees of quantity (Finnish 'kestoaste'). The Finno-Ugrian transcription system allows a division into eight different quantity degrees (Sovijärvi & Peltola 1964, p. 6–7):

over-short	Finnish	ylilyhyt
under-short		vajaalyhyt
short		lyhyt
half-short		puolilyhyt
half-long		puolipitkä
under-long		vajaapitkä
long		pitkä
over-long		ylipitkä

In the introductory booklet which presents the system it is, however, emphasized that these eight quantity degrees are only taken into consideration in a fine ('curve analytical') transcription; otherwise the following six quantity degrees are sufficient: over-short, undershort, short, half-long, long and over-long. This kind of system is hardly intended for the description of any functional linguistic unit called quantity but presumably for the subjective impression which the researcher has about the timing relationships of the sounds of the language form he is describing. What the researcher describes when using these quantity degrees is which of the sound segments of the language seem to differ from normal timing patterns in his own language, and how they differ. It is hardly possible to imagine that there could exist eight or even six absolute, distinctive classes, 'quantities', by which the normal listener should be able to identify and classify the durations of the vowels and consonants which he hears. Instead, one can imagine an appraisal in which the listener compares the phenomenon being described, for example his conception of the quantity of an inter-vocal consonant in the language being introduced, to that which he has of the corresponding structure of his own system. He can obviously observe that it is clearly shorter, slightly shorter, similar or slightly longer, etc., than his impression of the equivalent structure in his own language. This kind of evaluation of quantities is therefore substitution: the listener must have an idea of a 'short' and 'long' quantity in order to use his comparative scale. But there is no absolute, universal 'short' and 'long' quantity. The scales of substitution are the patterns of his own code.

For the naive Finnish speaker or listener there are two categories which he can call quantities, or simply 'long' and 'short': the sound segment of a word, which he writes with one letter, is short and the sound segment which he writes with two identical letters is long. Correspondingly he has the auditory impression that a sound written with two letters takes a longer time to produce than one which is conveyed in writing with one letter. For example, in the standard Finnish word, *mana* 'sorcery' both *a*-vowels are short (in the opinion of the Finnish listener), but in the word *maana* 'as a land' the first vowel is long and the latter vowel short. If the listener is asked what distinguishes the words *mana* and *maana* from each other, he confirms that in the former the first *a* is short, while in the latter it is long (unless he says that there are two *a*'s at the beginning of the latter). The Finnish listener can observe no other difference. The second vowel is short in both words. However, the second vowel of the *mana*, when measured, can be even more than twice as long in duration as the second vowel of the word *maana*.[1]

What an ordinary user of the language 'hears' in the quantity of a speech sound is only its function. He identifies the first vowel in the former example as belonging to either the word *mana*, or to the word *maana*. This categorization is always made whether the vowel is normal short, normal long or something else. The evaluation of subjective quantity depends on the category chosen by the listener. If a speech stimulus is used where the characteristics influencing the choice are half-way between normal or typical /maana/ and typical /mana/, the impression of phonetic quantity depends on whether the listener identifies the stimulus as the word *mana* or as the word *maana*. If he chooses the former, he describes the [a] of the first syllable as being longer than usual; (if he has linguistic training, he can imagine it as 'half-short' or 'half-long'). If he takes the stimulus to be the word *maana*, he thinks that the first vowel is shorter than usual, perhaps 'under-long'.

The reason why the quantity concept is so indefinite could at least partly be explained by the fact that several phenomena on different levels might have been added to it: on the one hand quantity may mean linguistic distinction, on the other hand it may mean the subjectively perceived dimension of time, or objective time measured in seconds and fractions of seconds. A general tendency noticeable in linguistic treatment of language has been to abandon the quantity-term as being indefinite in content, and to use the terms *length* (Finnish *pituus*) and *duration* (Finnish *kesto*) instead. Both these terms have more precise meaning and application. The duration of a phonetic segment means the time it actually lasts. One can also talk about duration on the various levels of production and reception: it is always simply the dimension of phenomena under consideration in time – it can have any function in the phonemic or phonetic system or in the perception of speech. Correspondingly, the term *length* is reserved for phonemically distinctive features or supra-segmental phonemes which seem, subjectively, to have something in common with observations of the production time of corresponding phonetic segments (see, for example, Fischer-Jørgensen 1964). But on the other hand it is sometimes given a wider significance: it can mean the time which a sound is heard to last regardless of the phonemic function of the phenomenon (Hadding-

[1] see chapter 3. 7.4.

Koch & Abramson 1964, p. 94). This latter interpretation also corresponds to the *subjective quantity* defined by Malmberg (1944, p. 28). Malmberg calls a linguistically relevant quantity the subjective quantity. This linguistically relevant quantity contains all those degrees of duration (Längestufen) 'die von der Aussprachenorm der betreffenden Sprache bedingt sind'. According to Malmberg's interpretation, the category of subjective quantity would not only contain phonemic quantity distinctions in the narrowest sense, 'sondern ausserdem auch jedes quantitative Verhältnis, das zum inneren Bau der Sprache gehört und von deren besonderen Gesetzen bestimmt ist'. One may however ask to what extent these durations are 'von der Aussprachenorm bedingt'. If by duration is meant only the dimension of separate phonetic phenomena in time, and by linguistically relevant everything associated with the durational patterns which act in perception as a cue of some distinction or linguistic category, almost all variations in temporal patterns of segments of varying size and all articulatory timing patterns can be seen as linguistic and therefore belong to the area of subjective quantity. In this area of quantity a great number of durational distinctions in various languages should be included which are not subjectively perceived as the time differences of speech sounds. It is apparent, however, that when one is talking about quantity or length, one is thinking in particular of those durational differences which are subjectively interpreted as a variation in the time used for producing the sound and which in various languages seem to be able to distinguish between otherwise identical expressions. In the present author's opinion the traditional concept of quantity should be divided into not two, but three distinct phenomena and possibly given three terms as well: in addition to objective, measurable duration and the phonemic length of the unit, it may be possible to speak about subjective or perceptual quantity – but not subjective in the sense defined by Malmberg. Here subjective quantity means that impression which a listener may have of the articulation time needed for some sound segment of an expression heard by him.

It is important to notice that duration, length and subjective impression of quantity are not in a one to one relation to each other: duration can serve as a phonetic parametre of many other phonemic distinctions besides the short/long distinction or the distinction of length.[1] Nor is it presupposed that the

[1] Duration can be one of the phonetic cues distinguishing different segmental phonemes, such as the duration of vowels before consonants in the distinction between voiced and voi-

shorter or longer duration of sounds would be the only or, from the point of view of perception, the most important phonetic cue of phonemic length distinction. As for subjective quantity, this is a very complex observation which is related to the real durations of the manifestations of phonemes and their contrastive ratios and to the identification of phonemic length.[1]

1.2. *Interference in physical characteristics, perceptual parametres and linguistic functions*.

Subjective observations of the quality (timbre), loudness and pitch of a voice are known to be based on the complexities of various physical characteristics: for example, observation of sound loudness apart from its amplitude, can also be influenced by its spectral pattern and pitch. But the perception of speech units differs significantly from that of non-speech stimuli. Linguistic 'hearing' is, in character, categorizing and selective: in a linguistic signal the listener perceives functions, not the colour, loudness, pitch and duration of the stimulus (see Malmberg 1963, p. 165).

Even when an attempt is made to evaluate linguistic stimuli audibly by common parametres of perception or when a listener (for example, a so-called 'ear' phonetician, German 'Hörphonetiker') tries to evaluate the absolute, subjective characteristics of phenomena, his observation is bound to the structure of the language and its functional units. It can be said that the evaluation of linguistic stimuli takes place as if it were on a different level from the perception of non-linguistic stimuli (Flanagan 1965, p. 231). When analyzing acoustic parametres of different linguistic functional units, the basis of all constructions must be the fact that what is under examination is a communication code between people, constructed according to the inherent potentials and limitations of the human being in the production and distinguishing of acoustic signals. These characteristics in the acoustic signal, which carry information of different linguistic functions, cannot always be described by means of one physical characteristic of voice (amplitude, fre-

celess consonants in English (c.f. Denes 1955) or oral vowel and nasal vowel distinction in French (Delattre & Monnot 1968). The duration of sounds can indicate a word boundary or some other phonological boundary, it can indicate stress, etc. In these cases the durational difference of the sound segment of the contrastive utterances is not perceived as a time difference or as a different quantity as is the case in the phonemic short/long distinction.

[1] For more discussion of subjective quantity, see Lehtonen 1969.

quency, wave-form, time), nor by means of one perceptual parametre. Apart from segmental sounds, this complexity is also true of suprasegmentals. The feature of speech known as intonation cannot be identified simply with variations in the oscillation frequency of the vocal chords, and its variations are not perceived in the same way as variations in non-speech tonal stimuli. The feature called stress, or accent, and subjectively experienced as greater 'weight' on a certain syllable of a word or on a certain word in a sentence (the linguistic term for *stress* in Finnish, *paino*, means 'weight', as in German and Swedish) is a typical 'complex quality', observation of which is based on a bundle of many acoustic and perceptual characteristics of speech signals (change in fundamental frequency, longer duration, rhythmical patterns and at times, possibly also intensity) – and on linguistic structure. However, if stress is phonetically manifested, it can obviously be described on the production level only by means of one articulation command (Lieberman 1966)[1]. Simple articulation of stress causes, or may cause, several different changes in an acoustic signal which are all parametres of stress on the acoustic level. The listener, however, does not perceive these separately; he does not 'hear' the change of melody, longer duration etc., but only that particular unit to which, according to his experience, these kinds of variations should be related: stress. One unit or one function in production or interpretation can, therefore, be a very complex phenomenon in a transmitting signal. Of these variations in the signal, caused by 'stress articulation' some (change of pitch, longer duration of the stressed syllable or of some sound segments in the word) can be shown to be more important cues of stress in perception than some others, but what the listener perceives is, however, stress or prominence and not a change in pitch or in duration.

The phonemic length is comparable to stress in certain respects: the listener perceives certain sounds as longer and certain sounds as shorter and we can presuppose that the speaker correspondingly intends certain sounds to be heard as long and certain ones as short and that he regulates the production of sound segments in such a way that a distinction between them will be

[1] In his later work Lieberman (1970, p. 314–315), however, emphasizes the potential complexity of the feature of *prominence* even at the articulatory level. Although the primary articulatory correlate of the state + *prominent* (i.e. phonetical manifestation of stress) is an increase in subglottal air pressure, it also may result in greater activity in virtually any of the muscles of the vocal tract. The perceptual effect of the complexity of the correlates is, however, merely that of »stress» or »prominence».

possible. But the characteristics of acoustic output by which the listener can estimate whether a speech sound is long or short can be various and complex, just as in the case of stress. If the quantity of sounds in a language has a phonemic function, in other words if different words in the language can be distinguished from each other only by some sound segment being heard to last longer in one word than in another, it must be presumed that the corresponding structures also include some difference on the acoustic level which make it possible for the listener to identify the word as one or other of the alternatives. There can be more and different acoustic characteristics making distinction possible than duration of the qualitative segment concerned: the quality of the segment, its intensity, the fundamental frequency during the segment, etc. The importance of these different phenomena as a cue to the identification of length can vary from language to language. The main point is that a language can make use of two distinctive categories, which it is possible to call, short and long on the basis of a subjective impression. The corresponding characteristic which distinguishes between them is called length. Those phonetic characteristics with which words can be distinguished from each other, a word where a certain segment is short from a word whose corresponding segment is perceived as long, are not necessarily limited only to that acoustic segment which mainly represents on the acoustic level the corresponding distinctive word structure segment. It can also naturally be supposed that in order to make the distinction, contrast between the successive segments of the word is also used within those limits set by the structure of the language.

Quantity or, more precisely, subjective quantity was defined as the time of articulation of a speech sound perceived by the speaker or by the listener, in other words as the listener's image of the duration. The linguistic function of the quantity and its phonemic interpretation can vary from language to language. In the same way, phonetic parametres causing a different image of the subjective quantity can vary from language to language. The term *quantity* naturally indicates that it would be possible to base perceived differences in quantity on the different extensions of sounds in time. All researchers have not, however, accepted that the duration of a speech sound is among the parametres contributing to the perception of its varying quantity. NS Trubetzkoy has even denied the validity of quantity conception in linguistic description. In his article, 'Die phonologischen Grundlagen der sogenannten Quantität' (1938) Trubetzkoy suggests that the structure of language (Sprachgebilde) as such is timeless (zeitlos). The concept of duration of speech sounds (Lautdauer) belongs only to the vocabulary of a field of science dealing with the problems of the standardization of the phonetic aspects of speech action (Sprechhandlung). But, according to Trubetzkoy, 'quantity' cannot be considered a distinctive feature of the phonemic system of a language either, since it is caused, in every language, by some primary distinction (in intensity, in tonal pitch, in 'Silbenschnitt' etc.). Therefore the terms 'Quantität', 'Länge', 'Kürze' should not be used at all in phonemic description: 'sie sollen nur als vorläufige Bezeichnungen dienen, d.i. nur dort, wo das phonologische Wesen der entsprechenden Erscheinungen noch nicht festgestellt worden ist oder aus Mangel an Material nicht festgestellt werden kann' (op. cit. p. 174). In basically the same spirit Marguerite Durand (1946) has examined the phonetic manifestation of long and short vowels. According to Durand the following characteristics can be at the root (à la base) of the phonemic distinction short vowel/long vowel (Durand 1946, p. 162):

'short vowel'	'long vowel'
– lax vowel	– tense vowel
– increasing intensity	– decreasing intensity
– increasing respiratory effort	– decreasing tension
– rising pitch	– decreasing respiratory effort
– separated from the subsequent consonant with a kind of close contact	– falling pitch
– possibly shorter duration	– separated from the subsequent consonant with a kind of open contact
	– possibly longer duration

It can be seen that Durand considers the longer duration of the long vowel the weakest characteristic distinguishing a long vowel from a short one: a long vowel may also be, in addition to other more important characteristics, longer than a short one! The most recent studies, performed for example by means of synthetic speech, have, however, also indicated the importance of the timing factor, as a parametre of phonemic length and of many distinctions not related to the quantity concept in perception (see e.g. Lenneberg 1967, p. 97–). The significance of the durational difference in the long/short distinction varies, as expected, from language to language, but also varies within the different vowel distinctions of the same language. It has been shown in Swedish that the duration of a vowel plays a decisive part in phonemic distinctions where the spectral difference between vowels is slight (for example, between the vowels of the words *väg* and *vägg*). On the other hand, in vowel pairs where the difference in quality between long and short vowels is clear (for example between the vowels of the words *ful* and *full*) a reduction of duration of the long vowel has very little effect on the listener's identification (Hadding-Koch & Abramson 1964). Similar results emerged from research on English and German long/short or tense/lax vowel distinctions (Bennet 1968). According to Bennet it would seem that in English and German spectral form is, in general, a more important cue for the tense/lax distinction than duration, but in distinctions where long and short vowels are similar in quality, duration is the decisive cue in both languages. Delattre and Hohenberg (1968) have examined in greater detail, by means of synthetic stimuli, the perceptual parametres of the tense/lax (=long/short) distinction in the German vowel pair ẹ/ę (=e:/e). Vowel duration appeared to be a very strong cue for the tense/lax distinction when vowel colour is held constant. But vowel colour alone also had a very strong effect on judgements. This means that the two cues contribute equally well to the tense/lax distinction in German. In addition to the quality of the vowel and its own duration, the duration of the consonant following the vowel was also shown to be an effective factor in the observation of vowel length: when the vowel was kept the same, extension of the consonant caused an increase in lax-identification.

In the Thai language, one of the so-called tone languages, the difference between long and short vowels would, according to Durand, be related to their different pitch-patterns. However, Abramson (1962) has pointed out that shortening or lengthening of vowel duration can alone contribute to the identification of a vowel as short or long (phonemically as a single or double vowel).

In this connection it may be worth examining the factors, other than durational difference, which also have a possible effect on the short vowel/long vowel distinction in Finnish. These factors may be compared to the parametres of Estonian quantities treated in the past in greater detail, even though the languages differ from each other in respect of distribution of quantities and their phonetic manifestations. In both languages the quality differences between shorter and longer vowels should be taken as an automatic concomitant of phonetic characteristics, which plays no decisive role in the perception of quantities (Estonian: Liiv 1962; Finnish see chapter 3. 7. 1.). On the other hand, it has been shown in connection with the Estonian and Finnish languages that the pitch pattern of a word or a syllable is different in different degrees of quantity. The differences in the pitch pattern of the Estonian language are obvious: a vowel with a first degree of quantity is characterized by a steep rise in fundamental frequency at the beginning of the vowel, a vowel with a second degree of quantity is characterized by a rise in fundamental frequency towards the end of the vowel and one with a third degree of quantity by a generally descending pitch pattern (Liiv 1961, pp. 488–489). It has even been suggested, that from the point of view of quantity perception the contrastive pitch patterns might be decisive and the durational differences might only play a secondary role in Estonian (Trubetzkoy 1939, p. 178; Durand 1946, p. 80–83). Liiv (1962, p. 285–286) has, however, shown by means of perception tests that different pitch patterns could not alone provide information about the degrees of quantity. The most important factor determining to which degree of quantity a vowel belongs, turned out to be the ratio of the vowel duration of a stressed syllable to the vowel duration of subsequent syllable.[1] Observations have also been made in connection with Finnish about the pitch patterns of short and long (phonemically single and double) vowels. Malmberg (1949, p. 43–45) has measured on kymograms, with the help of Meyer's pitch recorder, the pitch patterns of single and double vowels of a few Finnish words. He observed that all double vowels are characterized by a descending pitch pattern in the final part of the vowel, whereas the pitch pattern of short vowels rises at the beginning and then levels out.

Melody curves were made of all the sentences in the present study for min-

[1] The quantity system of Estonian has received a great deal of treatment in the light of experimental phonetic material and phonemics. Most Estonian researchers are of the opinion that in the first syllable of an Estonian word there is an opposition of three phonemic degrees of quantity and that the duration of the vowel of the second syllable is automatically determined in accordance with the length of the first syllable. However, the Finnish linguist Lauri Posti has interpreted Estonian quantity patterns within the framework of the binary short and long quantity distinction (see Posti 1968; Lehiste 1960, 1965 A, 1966 and references noted there). According to Posti's conclusion, long and over-long duration in the first syllable are only 'allochrones 'of the phonemically long quantity, the former occurring before the phonemically long and the latter before the phonemically short second syllable vowel. Apart from the observations concerning the necessity of the second syllable in the recognition of the quantity of the first syllable vowel, the fact that there are only two relevant lengths of vowel in a monosyllabic words in Estonian, short and long, supports Posti's conclusion.

gograms with Frøkjær-Jensen's Trans-Pitchmeter. These curves do not confirm Malmberg's results. In a tertiary stressed sentence position the pitch pattern of both long and short vowel segments is generally even, or more precisely even-descending owing to the descending sentence intonation. In the unstressed syllables after the first syllable of a word, the pitch patterns of both long and short vowel segments are identical. However, a rise in the pitch pattern at the beginning of a vowel segment appears in word-stressed syllables. The stronger the sentence stress falling on the word, the greater is the interval in the rise. In their duration these contours of the pitch pattern are of about equal value in normal utterances. If the vowel of a stressed syllable is short in its duration, the peak of tonal pitch may appear on the following consonant or if this consonant is short, not until the vowel of the second syllable. Of course, the rises in tonal pitch are caused by stress and as such are independent of the duration of those sound segments in whose area the suprasegmental feature of stress appears (for the parametres of stress, see Sovijärvi 1958). In the examination of pitch pattern, attention should also be paid to the effect of the characteristic descending intonation towards the end of the sentence in Finnish. Words which are pronounced in isolation receive the falling intonation pattern of a normal sentence, wherefore the pitch of a long word-final sound segment naturally descends (it also descends in a phonemically single half-long final vowel of the CVCV-structure), if, on the other hand, a final vowel is phonetically short, the characteristic descent in the tonal pitch for the terminal intonation occurs, in the main, during the previous vowel or consonant. In the present form of standard Finnish as well, obviously, as in large parts of the dialects of Central and Eastern Finland, variations in tonal pitch act as parametres of stress alongside a particular expressive function of the pitch pattern.[1] However, it is possible that in certain dialects of West and South-west Finland different pitch patterns act as potential indicators of various word structures which is something like what happens in Estonian.

A difference in contact of a vowel with the following consonant has been also suggested as one possible cue of quantity perception (see Liiv 1962, p. 268). This matter is treated in greater detail in connection with consonant quantities.

[1] Expressive modifications of the stress-parttern and intonation are described in Lehtonen 1968.

2. THE FUNCTION OF QUANTITY IN FINNISH

2.1. *The phonemic structure of Finnish*

2.1.1. Segmental phonemes

The Finnish phoneme paradigm includes eight vowel phonemes /i y u e ö o ä a/ and thirteen consonants /p t k s h m n ŋ l r j v d/.[1] The phonemes in this list are distinctive segmental and commutable entities which may be identified by certain contrastive articulatory or acoustic properties (distinctive features). This paradigm does not include phoneme entities which may possibly be superimposed upon sequences of segmental phonemes, or may influence the distribution of segmental phonemes (e.g. pitch, stress, juncture, measure boundary). Neither the position of these latter items in the phonemic structure of Finnish, nor their phonetic character have yet been systematically studied.[2]

[1] The vowel symbols *ö* and *ä* of Finnish orthography are used in this presentation instead of the corresponding IPA symbols (ø, and æ.) Other transcriptions used in the present work are in accordance with the IPA system if there is no remark to the contrary. In the phonemic transcription of Standard Finnish, letters differing from the orthography do not need to be used very often as the orthography is almost completely phonemic, at least on the level of word-phonology. As far as this presentation is concerned, the noteworthy differences between the orthographic and phonological forms are as follows: 1. The single velar nasal /ŋ/ which within a word only appears in front of the velar plosive /k/, is represented orthographically by the letter *n*, for example *kenkä* 'shoe', phonemically /keŋkä/. 2. The geminate nasal /ŋŋ/ is represented orthographically by the digraph *ng*, for example, *kangas* 'cloth'= /kaŋŋas/. 3. The final consonant in a certain group of morphemes, which has disappeared from its position before a pause but is manifested within the sentence as a gemination of the initial consonant of the following word (or as a doubling of the hard onset of the vowel or laryngalisation i. e. as a long glottal stop) is not visible in the orthography, for example, *tule tänne* 'come here', *vene upposi* 'the boat sank'= morphophonemically {tule* tänne*}, {vene* upposi}, phonemically (usually) /tulet tänne/, /veneʔ: upposi/. More about the relationship between phonemics and orthography in Finnish can be found in Sebeok 1944.

[2] The phonemics of Finnish is here treated according to traditional, so-called autonomous phonemics. (For the concept of autonomous or taxonomous phonemics, see, for example, Postal 1968, p. XI- and Karlson 1969, p. 351). The terms *phonemics* and *phonemical* are used in the present work as equivalents of the German terms *Phonologie, phonologisch* of the Prague School instead of more ambiguous terms *phonology* and *phonological*.

The arrangement of the vowels can be described as follows by means of articulatory criteria:

	front		back	
	unrounded	rounded	unrounded	rounded
high	i	y		u
mid	e	ö		o
low	ä		a	

As the table shows the system is based on three degrees of tongue height, on the articulatory feature front/back and on the rounding of lips. Lip rounding is a distinctive feature in the pair of high (=close) front vowels i/y and the mid (= half-close~half-open) front vowels e/ö, but not in the open vowel group. The distribution restriction of vowels called vowel harmony makes it possible to reduce the number of segmental vowel phonemes to five by making use of the suprasegmental phoneme /¨/, whose phonetic realisation in articulation is 'fronted tongue position' (Wiik 1965, p. 40). The application of suprasegmental /¨/ which simplifies the vowel paradigm as well as the definition of the distribution rules of vowels (and the morphophonemics), results in the breaking of the linearity principle of phonemisation and in the removal of the front/back feature of a phonemic segment away from its phonetically correct position in the phonemisation (e.g. the word *työ* 'work', according to traditional phonemisation /työ/, should be phonemised /¨tuo/; Karlsson 1969, p. 353). In the present study the traditional interpretation of eight vowel phonemes has been adopted in the phonemisations. Different kinds of models for the description of the Finnish vowel paradigm by acoustic criteria (on the basis of the F-patterns or distinctive features defined on the acoustic level) are introduced by e.g. Jakobson & Fant & Halle (1955, p. 41), Wiik (1965, p. 42–44) and Sovijärvi (1966, p. 191).

In the present list of consonant phonemes /d/ has been left last because its position among the consonant phonemes is marginal. It appears within a very narrow environment: /d/ appears only in word medial position between vowels as a single consonant. Morphophonemically, /d/ is not an independent unit at all, because it only appears as /t/'s weak equivalent in words partaking in the so-called gradation of consonants. However, on the phonemic level /d/ must be included among the independent units, because minimal pairs with the distinction t/d can be found (*katon* 'of the roof' /*kadon* 'of the lack' etc.). Sovijärvi (1966) has put on the paradigm of the Finnish consonants as many as 18 consonant phonemes, viz./k g t d p b š(=ʃ) s f ŋ n m ʔ φ r l ɦ h/; in addition to these Sovijärvi puts the semi-vowels /j v/[1] among the vowel phonemes in his branched-tree diagram showing contrastive hierarchic relationships. According to the present writer, the phones [b g ʃ f] which, in fact, can appear in educated speech in the pronunciation of unabsorbed new loan words in Finnish should, at least for the time being, be excluded from the system of Finnish distinctive phonemes (see also Hakulinen 1961, p. 22–23). Also the bilabial trill [φ], which at least serves as a signal for stopping a horse (!) should not be considered a linguistic unit. The phonemic distinction h/ɦ should, according to Sovijärvi's example, hold true only at the boundary of a word *puu+hanko* 'a wooden pitchfork' [pu:hanko], *puuhan+ko* 'the job's?' [pu:ɦanko]. The voicelessness of /h/, defined by Sovijärvi would thus be one of the many phonetic manifestations of internal juncture, along with the glottal plosive (Lehiste 1964, p. 177, 181). The voiced [ɦ] is, in any case, the most common allophone of /h/. /h/ is quite clearly voiceless except at the beginning of an utterance when it appears as the initial member of a consonant cluster. However, the following are also fully acceptable and common allophones of /h/: the velar fricative [x] and even palatal [ç] when /h/ appears with high vowels or /j/, for example *vihje* 'cue' [viçje], *tuhka* 'ashes' [tuxka] (cf. Wiik 1969, p. 227). Observations about the distribution and the articulatory phonetic classification of the consonant phonemes can be found in Karlsson 1969.

2.1.2. Some phonotactic restrictions

It may be necessary to deal in this connection with those restriction rules of the distribution of Finnish phonemes which are concerned with the appearance of phonemic units, whose phonetic manifestation is subjectively interpreted as long duration of a vowel or consonant. Here 'long quantity' is phonemized as a sequence of two identical segmental phonemes. This solution and other possibilities for the phonemic interpretation of quantity will be examined in greater detail later.

[1] Wiik (1969) has made different suggestions for the phonemisation of Finnish *j* and *v* based, on the one hand, on generative phonology, and on the other hand, on autonomous phonemics.

The frame of reference, within which the distributional rules function, is a phonemic word, a phoneme sequence, which functions as a free morpheme and which can include a varying amount of bound morphemes (endings etc.). A Finnish word always includes at least two phonemes of which at least one must be a vowel. The simplest word structures are therefore CV (e.g. *te* 'you, pl.') VC (e.g. *on* 'is') and VV (e.g. *yö* 'night'). There are no words in the language consisting of only one vowel phoneme, nor are there any VV words in which both vowels are the same. Longer words are divided into syllables by means of fairly simple rules. Every syllable includes as a syllabic either one vowel or a sequence of two vowels. A syllabic vowel sequence can be made up from a combination of two similar vowels or any of the following combinations of dissimilar sounds *uo yö ie ai oi ui ei äi öi yi au ou eu iu äy öy* and in some words also *ey iy* (Penttilä 1963, p. 19). Other kinds of vowel combinations contain (at least potentially) a syllable boundary. A word can begin and end with either a vowel or a single consonant. Consonant clusters do not appear at the beginning or at the end of a word. A single consonant or a combination of not more than three consonants, of which the latter members must be some of the obstruents /ptks/, can appear within a word between vowels. The syllable boundary is placed before the final consonant of a consonant combination within a word (cf. *pa-ta* 'pot', *pas-ta* 'paste', *pals-ta* '(wood)lot'; *ka-sa* 'heap', *kan-sa* 'people', *kans-sa* 'with'). Because only the last two members are allowed to be alike in a sequence of three consonants, the sequence of two identical consonant phonemes everywhere automatically includes the boundary of the syllable. Of the consonant phonemes in intervocal position, all except /j v h/ can appear as geminates. Of the three exceptions, /j v/ cannot appear as initial members of an intervocal consonant cluster either. The plosive phoneme /d/ appears as a geminate only in new unabsorbed loan words (cf Swadesh 1937, p. 3, 5). Double vowels and double consonants appearing within a word are independent of each other; a single vowel can be followed by either a single or double consonant. In the same way a double vowel can be followed by either a single or double consonant. However, a cluster of three consonants cannot appear after a double vowel. The appearance of vowel combinations[1] (i.e. combinations of two similar sounds, monophthongs and most diphthongs) in the same syllable is independent of the length of surrounding consonants, and also of stress and the position of the vowel segment in a word.

[1] The diphthongs *uo yö* and *ie* appear, however, in stressed syllables only.

2.1.3. Phonemic word structures

Syntagmatic phonemic units smaller than words are generally said to be syllables. Two syllable quantities may be discerned in Finnish: short and long. Short syllables are those which end in a single vowel, the others being long. Syllable-initial consonantism does not effect the quantity of the syllable (Sadeniemi 1949, p. 101).

That the flow of speech is formed into syllables is a phenomenon whose role in the structure of language cannot be denied. The concept of syllable units is also useful in the treatment of many of the problems in Finnish phonology, even though it is not generally necessary since the boundaries of syllables are automatically fixed in front of each CV sequence. In the phonetic treatment of durations of sound segments the concept of the syllable cannot, however, always be used when a long consonant segment contains the boundary of a syllable. The phoneme sequence /VCCV/ includes the syllable boundary between the consonant phonemes, but when the consonants between the vowels are identical, there is phonetically only one consonant segment between the vowels [VC:V]; the arrangement of the duration of this segment between the preceding and following syllable cannot be indicated phonetically. For this practical reason, the word structures are described in the present work as sequences of vowel and consonant segments without any marked division of the grouping of segments into syllable units in each separate structure. The concepts of vowels and consonants are employed here to indicate the functions of phonetic sound groups: a vowel segment is that phonetic sound segment which acts as a syllabic in a word; a consonant segment is that phonetic sound segment which functions as a non-syllabic. In practice, therefore, the number of syllables is the same as the number of vowel segments in the word structure.[1]

The position of the sound segment in the word structure is indicated by a small numeral after the vowel sign (V) and after the consonant sign (C). The vowel segments (syllabics) of a word have the following numerals: V_1 V_2 V_3 etc. and the consonant segments C_1 C_2 C_3 etc. These signs, however, only indicate the position of vowel or consonant segments in a word; a V-segment

[1] Words containing so-called hiatus-sequences (e.g. *teossa* 'at work', phonemically /te-ossa/) are not included in the material of analysis of the present work. Description and analysis of the hiatus-sequences compared to proper diphthongs will be found in Karlsson 1970. With respect to the temporal pattern of a word the effect of the boundary marker /–/ is analogical with single consonants.

with some of these position-indicating numerals can be phonemically a single or a double phoneme (or a diphthong) and a consonant segment (C) between sonants can be a single consonant or a sequence of (identical or different) consonant phonemes. The initial consonant of a word is marked C_1, the consonant between the first and second vowel C_2 etc. The following examples illustrate the system of symbols used:

phonemic word structure	example	position			
		C_1	V_1	C_2	V_2
CVCV	*taka* 'back'	t	a	k	a
CVVCCV	*taakka* 'burden'	t	aa	kk	a
CVCCV	*taksa* 'fee'	t	a	ks	a
CVVCV	*taika* 'magic'	t	ai	k	a

As is apparent from the preceding, with regard to word structures, a phonemically double V segment can be formed by two identical or different vowels, just as a phonemically double C segment can be a sequence formed by two identical or different consonants. However, in the present material for the analysis of the durations, all /VV/s are combinations of two identical vowels and not diphthongs. /CC/ segments are, in the present analysis material, geminate consonants, unless it is indicated by a certain sign (C_XC_Y) that in the words representing the word structure under consideration the consonants between the vowels are different.

2.1.4. The number of word structures

The number of phoneme combinations appearing as words in Finnish, or which are possible according to Finnish phonotax, is rather great. It is hardly possible to analyse all of the combinations of various vowel and consonant segments of different length appearing in the language. But even when we disregard the phonemic quality of consonant and vowel segment, and examine the possible consonant and vowel segment combinations in Finnish words of different length, we find ourselves dealing with large numbers. According to Lehiste's formula (1965, p. 448):

$$P_n = 2^{2n+1} \text{ (n= number of syllables)}$$

the number of combinations of single and double vowels and single and double consonants is 32 in two-syllabled words, 128 in three-syllabled words and 512 in four-syllabled words. It is supposed in this formula that a word can begin with a vowel or a consonant and end in a vowel or single consonant, and that each vowel segment appearing as a syllabic can be phonemically single or double. The absence or presence of word-initial consonants does not however affect the phonemic combination possibilities of other sound segments. When the influence of the initial consonant is omitted from the calculation, the number of structures is to some extent reduced: the number of two-syllabled structures where each vowel segment can be either a single or a double vowel and the intermediary consonant segment either a single or double consonant in function is 16, three-syllabled 64 and four-syllabled 256. Examples of all the possible two-syllable word structures are included in the material for this study: CVCV, CVVCV, CVCVV, CVVCVV, CVCCV, CVCCVV, CVVCCV, CVVCCVV, CVCVC, CVVCVC, CVCVVC, CVVCVVC, CVCCVC, CVCCVVC, CVVCCVC, CVVCCVVC. Out of the great number of possible three- and four-syllabled structures only a few combinations are included in the material. These longer structures are partly arbitrarily selected from those from which examples comparable to the words of the list of bisyllabics were found.

2.2. *Phonemic interpretation of quantities*[1]

In a phonemic analysis of languages containing a distinction between 'long' and 'short' vowels, that is where minimal contrasts between words seem to be based solely on a relatively longer or shorter duration of certain qualitative sound segments, there are various ways in which 'distinctive quantity' may be phonemically interpreted:

1) long vowels and short vowels are treated as separate segmental phonemes, in which the features long/short, tense/lax or some other corresponding feature may be taken as distinctive.

[1] For the general principles of the phonemic interpretation of vowel quantity, see, for example, Trubetzkoy 1938, p 169; Malmberg 1944; Pike 1947, p. 138; Jones 1950, p. 108-; Hockett 1955, p. 76–; Enkvist 1962; Elert 1965, p. 39–.

2) long vowels may be taken as a combination of a short vowel and the suprasegmental length phoneme /:/ or chroneme.
3) long vowels may be interpreted as a combination of a single vowel and some neutral consonant phoneme.
4) long vowels may be regarded as a coordinate construction of two identical short vowels.

In descriptions of Finnish where the sound system of the language has also been considered, the fact that Finnish vowels have two quantity degrees (short and long) is generally mentioned without any definition of the position of this distinction within the phonemic system.[1] The traditional and phonemically unspecified attitude of students of the Finnish language would seem to be analogous to the theory of Daniel Jones (1950). Jones draws a distinction between 'true' or 'indivisible' length and repeated or double sounds. The difference between double and single long vowels may be inappreciable to the ear, and may be even objectively non-existent to listeners, but whereas subjectively double long vowels may always be pronounced double, this is not the case with true long vowels. If a certain long vowel is never separable into two and cannot be replaced by a sequence of two in precise speech, then the case must be considered to be one of true length (Jones 1950, p. 116). The distinctive categories of true lengths are called chronemes. Since the long vowels of Finnish cannot be pronounced separately (as can the second vowel segment of the English word [emptiiŋ] in Jones's example), e.g. *ku-u (= *kuu*, 'moon') is not possible, the long durations of vowels should be interpreted as chrones of a long chroneme. But diphthongs, which are equal to long vowels in concepts of subjective quantity, could not be produced in speech as separate parts either. Therefore even these, according to this criterion, would be monophonematic long vowels (cf Karlsson 1969, p. 355–356). Since there is no long diphthong/short diphthong distinction in Finnish, the chroneme interpretation cannot be applied in the description of diphthongs. Long word-medial consonants can always, owing to the syllable boundary, be pronounced as two separate parts. Accordingly they should be interpreted, and have always been interpreted, as double sounds. In the Finno-Ugrian transcription system, too, (Sovijärvi & Peltola 1964) long vowels are represented by one vowel symbol

[1] See, for example, Hakulinen 1961, p. 24; Penttilä 1963, p. 21. Sovijärvi also omits this kind of definition even though he makes it clear that the opposition short/long is *prosodic* (Sovijärvi 1966, p. 190).

to which a diacritical mark has been added, but long consonants are represented by two letters. In Finnish orthography both subjectively long vowels and long consonants are represented by two letters.[1]

Using morphophonematic criteria, Trubetzkoy, in his 'Grundzüge', (1967 p. 170) interprets the long syllabics of Finnish as double vowels: 'Sicher polyphonematisch (d.i. als verdoppelte, geminierte Silbenträger) müssen die langen Silbenträger in solchen Sprachen gewertet werden, wo zwischen Anfang und Ende eines solchen Silbenträngers eine morphologische Grenze fallen kann (as in Finnish *talo-a*, *kukka-a*, *leipä-ä*, *kylä-än*). – – In allen solchen Fällen müssen die langen Vokale als die Summe von zwei gleichen kurzen Vokalen gewertet werden und diese Wertung darf dann auch auf alle lange Vokale derselben Sprache übertragen werden.' However, Malmberg (1944, p. 87) has challenged Trubetzkoy's interpretation. According to him the gemination of Finnish vowels and consonants should be treated synthetically and not analytically as Trubetzkoy did. The reason for this is that, for example, in the words *tuuli* 'wind', *kuu* 'moon' and *juusto* 'cheese' the long vowel belongs to the same stem and could not, therefore, be understood as a combination of two vowels thrown together.

Trubetzkoy's 'polyphonematic' view has recently received some favourable appreciations (Harms 1960, p. 7; Enkvist 1962; Wiik 1965, p. 41–42; Pilch 1968, p. 91, see however p. 73; Terho Itkonen 1968, p. 95; Karlsson 1969!). It is possible to refer to a number of criteria which support the pre-eminence of the analytical solution: long vowels are structurally equivalent to clusters of different vowels within the same syllable (see Pike 1947, p. 138); both the beginning and end of a long vowel are ceteris paribus commutable into some different vowel segment, the latter part is even commutable into a consonant, while the subjective syllable quantity stays the same i.e. long (cf word pairs *viinaa* 'of spirits' /*vainaa* 'dead'; *tuuli* 'wind' /*tuoli* 'chair'; *tuuli* /*tulli* 'customs' etc.). It has been pointed out earlier that the application of suprasegmental

[1] It was not until as late as the beginning of the 19 th century that two letters became established in the orthography as the only spelling for double phonemes (Rapola 1965, p. 69–76). In the past, in so-called old literary Finnish the representation of double vowels and of single and double consonants was uncertain: it was possible to represent double vowels, monophthongs and also diphthongs with one letter, double consonants could be shown by one and single consonants by two letters. However, there seems to have been, ever since the first preserved written texts in Finnish, representations of 'long' sounds with two letters according to present practice. In the 17th century there was, already, a marked consistency in the spelling of both vowel and consonant segments.

length phoneme is possible only for long vowels but not for long consonants. However, apart from the inclusion of one new paradigmatic phoneme unit, this would also mean that the syntagmatic combination rules would become more complicated. But when long vowels are interpreted as sequences of two identical vowel phonemes belonging to the same syllable, equivalent with diphthongs in the word structure, and long or geminate consonants are similarly interpreted as sequences of two identical consonant phonemes corresponding in the word structure to sequences of two different consonants, the description of structures becomes fairly simple. The following essentially phonetic facts may also be noted: so-called long vowels are about twice as long in duration as short vowels in the same position in a word; similarly, long consonants are twice as long as short ones. When the different intrinsic durations of sounds are taken into consideration, it may be said that diphthongs and long vowels, as well as geminate consonants and clusters of different consonants are equal in their phonetic duration. Furthermore, the effect of geminate consonants and clusters of different consonants and of both long vowels and diphthongs on the duration of other segments of the word structure is the same.

However, it cannot be stated categorically that either one of the two solutions, the use of chronemes or the interpretation of two identical phonemes, is the right solution and that the other one is definitely at fault. The choice of interpretation is finally a primarily theoretical question, depending on the demands of the phoneme inventory. It can even be said in connection with Finnish, that the interpretation of a long unit as a combination of segmental phoneme and segmental phoneme (which is manifested phonetically as the lengthened duration of a certain quality segment) and its interpretation as a combination of segmental phoneme and co-phoneme or chroneme (which is manifested phonetically as lengthened duration of a certain quality segment) are the same solution in practice, because the long quantity is, in both cases, interpreted as a combination of phonemes. However, as far as the needs of the present study are concerned, it may be practical to select the simpler of the interpretations, that which makes possible identical phonematic description of the quantities of vowels and of consonants.[1] Long consonantal quantity which,

[1] Phonemisation as double vowels or double consonants does not make the pair of concepts 'phonemically long'/'phonemically short' vowel or consonant necessarily mistaken. It is of use in particular in the treatment of the phonetic durations of sound segments. On the

in its distinctive function or in the phonetic realization of distinction, does not differ from long vowel quantity in any way, cannot be described by means of a chroneme: it cannot, for example, be stated that the boundary of a syllable in a long consonant should be fixed between a consonant phoneme and a chroneme. But the chroneme is also impractical in the treatment of vowels when we come to questions outside autonomous phonemics (for example, when explaining morphophonemic variation *puu/puita*, *nuo/noita* etc.). One possibility in the description of quantities would be to employ the quantity unit *mora* and to interpret the short sound as having a value of one mora and the long sound as having two moras. This would also comply with the biphonematic interpretation of a long consonant and long vowel. However, the mora should not be identified with the segmental phonemes, in such a way that the mora-number of the utterance is the same as the number of its phoneme segments.[1] Word initial consonantism does not effect the mora number, but neither does a syllable-initial consonant have any effect on the mora number of a word or on the rhythmic length value of the syllable (Sadeniemi 1949, p. 103; Lehiste 1960, p. 54, 76). The quantity structure of the words *kelossa* 'in a dried-up tree' and *elossa* 'alive' is the same, but also the word *keossa*/ke-ossa/ 'in a heap' is prosodically identical to the words *kelossa* and *elossa*. In all these words the total mora number is four. When syllable-initial consonants have no mora value, the syllable units are in fact unnecessary in the definition of mora structures. The words *koossa* 'together', *kuussa* 'on the moon', *tuossa* 'there' etc. are equal in mora-number to the words of the *kelossa*, *keossa* type. In this estimation of mora-numbers, all phoneme segments

phonetic level, the physically long duration of a segment in a known position could have some effect, for example, on the duration of the previous segment, even if the phonetically long sound represents a double or single phoneme: Such a phonetically long sound is said to be 'phonemically long' when it represents the combination of two segmental phonemes; in other words duration is phonemically long which serves to indicate the phonetic segment as a combination of two segmental phonemes. On the other hand, the concepts 'long allophone' and 'short allophone' of vowels should not be identified with single and double phonemes. If the terms 'long and short allophones' are used they could mean, for instance, those vowel qualities appearing in a sequence of two identical vowels and the allophones of vowel phonemes appearing in other environments. This division also seems to be unnecessary because of the insignificance of the difference.

[1] Itkonen (1968, p. 96) has defined the mora as the prosodic segment equivalent to a short single vowel (and consonant). According to this the constructions *apu*, *rapu*, *krapu*, 'prosodically' similar in rhythmic structure, would be quite different in their mora-number (cf Sadeniemi 1949, p. 102). For more about the mora concept, see also Trubetzkoy 1967, p. 164.

are equivalent to one mora except a consonant immediately preceding a vowel: for example, the words *ui* '(he) is swimming', *pui* '(he) is threshing', *puri* '(he) bit' all have two moras, while the words *kaato* 'overturning', *kanto* 'stump' and *katto* 'roof' have three.[1]

NS Trubetzkoy (1938, p. 158) has compared phonemically short and phonemically long quantity to point and line: 'Die Kürze ist punkthaft, indimensional, undehnbar; die Länge linienhaft, dimensional, beliebig dehnungsfähig.' Elasticity, Dehnbarkeit, is, according to Trubetzkoy, the correlation characteristic of a quantity correlation: long quantity is 'merkmaltragend' and short 'merkmallos'. The 'elasticity distinction' between long and short sounds is, according to Penttilä (1963, p. 21), essential in Finnish: short vowels or consonants cannot be noticeably lengthened or stretched whereas long vowels and consonants allow their duration to be considerably extended without any inconvenience. However, this is true only on certain conditions: the lengthening or stretching capabilities of sound segments depend on the position. The duration of a phonemically single phoneme can be freely lengthened in all positions where a sequence of two identical phonemes could not occur. In positions where the appearance of either a single or double phoneme is possible, the duration of the phonetical segment representing the single phoneme can only be varied within very strict limits, whereas the corresponding double phoneme can be lengthened just as one likes. Naturally, in the case of double vowels and double consonants it is not possible to say whether it is the beginning or end of the geminate which stretches, but if these are compared with parallel structures consisting of different segments, it can, at least in the case of the consonant cluster, be said that the first member stretches. For example, in the pronunciation of the isolated word *maltan* 'I am patient' it is possible, apart from the initial and final consonants of the word, to lengthen the first member of the intervocal consonant cluster [mal:tan] without changing the meaning of the word or making it sound unnatural. Similarly, the lengthening of the first two members of the intervocal cluster in the word *murskan* 'of crushed' is quite possible ([mur:skan murs:kan mur:s:kan] are all possible) whereas lengthening of the last consonant of the cluster is considered odd or wrong even though it does not effect recognition. Diphthong stretching is

[1] A mora here is not a unit indicating phonetic duration, but the mora numbers presented here also have a certain correspondence with the manifested phonetic durations. See chapter 3. 12. 2.

more complex. Lengthening of both the first and last member of the diphthong is possible in, for example, *laiva* 'boat', which can be pronounced [la:iva] or [lai:va]. Probably the information a listener receives from lengthening of separate segments is, however, slightly different: lengthening of the first segment may be more closely related to an expression of the speaker's emotional attitude (as in [ka:uheam pa:ljon] /kauhean paljon/ 'terribly much'), while longer duration of the second member is only taken as an indicator of slower speech tempo. In any case, the assertion that a double phoneme can be stretched while a single phoneme, point-like, cannot, is only partly true: double vowels and double consonants stretch but so do single word-initial and word-final consonants and consonants which are the first members of consonant clusters, as well as, obviously, single vowels which are the final members of vowel combinations.

3. THE ANALYSIS OF PHONETIC DURATIONS

3.1. *Concept of standard lanquage and the geographical position of Jyväskylä town*

The language form, whose sound durations and their linguistic functions this research endeavours to describe, is standard Finnish as used by the educated inhabitants of Jyväskylä. By standard Finnish we mean that particular language form which is used in the mass information media, such as on television, radio and in written form in the press and in official literature, and which is the aim of mother tongue education in schools. The literary language on which the original linguistic material of this research is based differs to some extent, morphologically and syntactically, from the colloquial language used by educated people (see Itkonen 1966, p. 9–), but does not differ phonemically or phonetically in a consistent way from this colloquial language. There are some noticeable differences in the spoken language even when the speakers endeavour to attain the normative forms of the literary language. The differences are perhaps clearer in the prosodic features of speech: in intonation, in phonetic manifestation of stress and in the quantity pattern of certain word structures. The latter difference has been clearly indicated by Wiik and Lehiste (1968) in their analysis of the speech of university students from different parts of Finland. If the small geographical phonetic differences are taken into account a homogeneous spoken standard Finnish does not exist: the spoken language of the educated class in Turku in the South West dialect area, in Kuopio, in the area of the Savo dialect and Helsinki, which comes under the influence of the South Häme dialect, includes in each case special phonetic features caused by the geographical local dialect. However, the thorough analysis of the differences in the educated standard Finnish spoken in various parts of the country is a task which, at the present stage, has been left for the future.[1] The present analysis, based on a geographically clearly limited material, is intended to be a basis for later comparisons. In the opinion of the present author, however, the regional differences in spoken standard Finnish will hardly change the results of this analysis concerning the general character of Finnish quantity patterns.

The town of Jyväskylä, where all the informants in this research reside, is in an advantageous position from the point of view of dialect geography, because it does not clearly lie inside the kernel area of any one main dialect. The parishes to the north of the town are in the area of Central Finland belonging to the transitional dialects of Savo, to the south of the town commence the dialects of Päijät-Häme, which can be considered a transitional area between the Savo and the Central Häme dialects. On the Southwest side of Jyväskylä there is the transitional dialect area between Central Finland, Upper Satakunta and Central Häme (see Itkonen 1964, pp. 30–31). Preliminary information about a study on the durational patterning of Finnish dialects (Wiik 1970; see the footnote on page 136) also indicates that the Jyväskylä area is near the boundary of dialects which are of different types in their durational characteristics: on the north and east side of the town the so-called half-long vowel appears in the second syllable of (C)VCV(C) structure words, which has not, however, been observed in the dialect area to the southwest of Jyväskylä. Irrespective of the vicinity of the borders of the transitional dialects between the eastern and western main areas, the town itself, according to the recent observations of Wiik, seems to fall inside the eastern domain of the half-long vowel. As the following analysis will show, the half-long vowel, or the durational ratio $V_1 < V_2$ in the /VCV/ strings, also exists in the speech of the informants of the present work.

3.2. *Material for the analysis*

3.2.1. Selection of informants

In phonetic research where the aim is a description of the manner of speaking of some geographic or social dialect or, as here, of the standard language,

[1] One interesting problem to be investigated is, how the educated speech with and without a half-long vowel (and the concomitant features) is stilistically evaluated by scholars of Finnish language or educated people with and without the half-long vowel in their own speech. It is possible that the listener prefers that form of timing corresponding to his own patterns and considers the other one to have a slight dialectal accent.

the selection of speakers is of fundamental importance. Success or failure in selection has a decisive effect on the interpretation and usefulness of results in the whole research.

The analysis results of this research are based on the measurement of the sound duration of words in sentences read by ten informants. The intention in the selection of readers was that they should represent the form of standard Finnish used by educated people whose speech is affected by analogous dialectal backgrounds and is, in the opinion of the writer, as neutral as possible.

All the readers are from about the same age group and have about the same educational, regional and social background. The selection was made solely on the basis of background information, partly with the help of the list of students at Jyväskylä University, partly by interviewing students. Speaker No. 1, who also confirmed beforehand, during his interview, that these persons did not have any striking pathological speech or voice defects, assisted in the latter task. No other criterion, either linguistic or based on speech habits, was used in the selection. Before the selection, the researcher did not judge the manner of speech of the persons himself according to his own subjective idea of standard Finnish or of correct pronunciation.

Within the domain of the Finnish language one can clearly observe that an urban area tends to reduce the dialectal characteristics of speech. Another factor having a similar effect is, naturally, education: one of the aims of education in the Finnish school system is the teaching of correct standard Finnish in pronunciation. It can also be seen that people of school age consciously endeavour to avoid using their own local dialect in colloquial speech. Instead of their local dialect, schoolchildren – at least in Jyväskylä – try partly to imitate the slang of the capital in their so-called grammar-school slang, but this kind of imitation is no longer noticeable among university students. The university community itself is also likely to level out dialect differences because a large number of students are from areas whose local dialect is different from that used in the Jyväskylä area. In the background factors of the selected speakers one should also take into consideration the possible effect of their parents' manner of speaking. According to the principles of this research, an ideal speaker would be someone who is an academically educated native of Jyväskylä and whose parents are also natives and have had as wide an education as possible. However, it turned out to be difficult to find an adequate number of readers from the rather small number of Jyväskylä inhabitants (total population less than 60,000) available in the university and complying with all these conditions.

All the readers were born in Jyväskylä town, had been to elementary school in the same town and had attended various grammar schools there. Each has studied 2–4 years in Jyväskylä University. There are five female students and five men. Next the information about their parents' origins and the year the parents moved to Jyväskylä is given. The schooling of the parents is also given, if it is other than compulsory elementary education.

Speaker Nr 1 (P. L.), male, born in 1947, student of Finnish language. Father born in Hankasalmi, in Jyväskylä since 1943, commercial school. Mother born in Jyväskylä rural commune, commercial school.

Speaker Nr 2 (E. F.), male, born in 1947, student of the history of art. Father born in Jyväskylä, business college. Mother born in Jyväskylä, secondary school.

Speaker Nr 3 (T. Y–P.), male, born in 1946, student of political science. Father born in Jyväskylä. Mother born in Korpilahti, in Jyväskylä since 1925.

Speaker Nr 4 (K. K.), male, born in 1944, student of sociology. Father born in Hankasalmi, in Jyväskylä since 1918, school for technical education. Mother born in Jyväskylä.

Speaker Nr 5 (R. P.), male, born in 1945, student of social politics. Father born in Keuruu, in Jyväskylä since 1912. Mother born in Korpilahti, in Jyväskylä since 1912.

Speaker Nr 6 (R. K.), female, born in 1946, student of speech therapy. Father born in Äänekoski, in Jyväskylä since 1910. Mother born in Jyväskylä.

Speaker Nr 7 (L. H.), female, born in 1946, student of English philology. Father born in Savonlinna, in Jyväskylä since 1923. Mother born in Jyväskylä, secondary school.

Speaker Nr 8 (H. T.), female, born in 1947, student of economics. Father born in Jyväskylä, business college. Mother born in Lapinjärvi, in Jyväskylä since 1930, nursing school.

Speaker Nr 9 (T. N.), female, born in 1946, student of Roman philology. Father born in Viipuri, in Jyväskylä since 1937, secondary school. Mother born in Jyväskylä, secondary school.

Speaker Nr 10 (P-L. S.), female, born in 1945, student of English philology. Father born in Isokyrö, in Jyväskylä since 1938, secondary school. Mother born in Jyväskylä, secondary school.

The voice characteristics and articulation of all the speakers turned out to be normal and free from defects. According to the subjective evaluation of the researcher, the manner of speech of all readers was perfect standard Finnish, in which no trace of dialect or slang could be heard.

3.2.2. Reading of test sentences

Each of the sentences to be read had been typed on a separate card. The reader had the cards upside down in front of him in a pile on the studio table or in his hand. From this pile he picked up one card after another and read the sentence from it, placing the card on another pile. This kind of system prevented the readers from joining the sentences together in a longer coherent sequence and also effectively regulated the speech-speed at the same time. During the reading each reader was given several short breaks (after approx. 50 sentences) and one longer coffee break half-way through. Every other reader was given the cards in the reverse order, and in addition the order of the whole pile of cards was changed so that each sentence was read at the beginning, end and in the middle of the series.

At the beginning of the test the readers were advised to read each sentence at a normal tempo and in a matter-of-fact voice without any 'artistic embellishments'. They were not told the purpose of the test. The reading of the whole material took each student about 2 hours. The readers were paid for their work.

3.2.3. Durational differences between speakers

As an indication of the individual differences and the differences between the sexes there is in table 1 a comparison of the durations of the sound segments and respective frame sentences of 22 CVCCVV-type test words as read by separate speakers. Since the individual variation has been taken into account in comparing the durations of separate linguistic units, no wider material is included to indicate individual differences. The significance of the differences of individual durations has not been tested separately for each speaker, but by means of a oneway variance analysis only the group as a whole (10 people), both women and men, was tested for the significance of duration differences of variables (i.e. sound segments).

If the means of the durations of the sentences are taken as a measure of speech tempo, speaker No 10 clearly has the slowest tempo. The sentence duration for speakers No 9, 8, 6, 3, 5 are slightly shorter and speakers 2, 4, 7 and 1 had the shortest sentence duration. The order would be approximately the same if the duration of a measured word were taken as a measure of speech tempo. On the other hand, there is a more irregular variation in the duration of different sound segments with different speakers. Even though the difference between the segmental durations of different speakers is statistically significant, the individual differences are, on average, relatively small. According to the durations of the sentences, the tempo of the quickest speakers (nos. 1 and 7) is only 9 % quicker than the average and the tempo of the slowest speaker (No 10) only 12 % slower than the average.

The differences between the sexes are not, on average, significant. The only statistically significant difference in the comparison is the duration of the single vowel of the first syllable, which is longer on the 5 % significance level with female readers than with male readers. This result does not accord with the observations made by Elert concerning the sound durations of four female and four male Swedish readers. According to Elert (1965, p. 179) the duration of both vowels and consonants is longer in the words read by male readers than in those read by female readers. The reading tempo of sentences is correspondingly 2.3 % quicker with women than with men.

In the present material this relationship also appears to be reversed: amongst five slower readers there is only one man. A superficial comparison of various test word series performed by the writer lends support, however, to the relationships indicated in Table 1. The average durations, in all word structures, seemed to be longer with female readers than with males. However, it should be noticed that the individual differences are more significant than differences between the sexes. The only conclusion concerning the sexes which could be drawn from results with this small test group is that, here, the tempo of speech and the duration of sounds do not correlate significantly with the sex of the speaker.

TABLE 1. Individual differences and differences between sexes in the segmental durations of CVCCVV-type words and in the durations of the corresponding frame sentences.

Speaker	Sound segment duration (msec)				Sentence duration
Nr	C	V	CC	VV	(msec)
1	66	69	181	133	1350
2	61	69	199	134	1430
3	69	87	212	125	1520
4	60	82	168	118	1360
5	60	64	204	143	1490
Men $\overline{X}$	63	74	193	131	1430
6	65	73	181	137	1520
7	52	71	168	112	1350
8	68	80	193	150	1520
9	72	89	178	137	1550
10	85	83	213	129	1660
Women $\overline{X}$	68	79	186	133	1520
Total mean	66	77	190	132	1475

3.2.4. Linguistic material

All test words used in this research as material for the analysis of sound durations are read within the frame sentences. On the other hand, the words used as material for the segmentation tests have been chosen from the inner words within lists containing 3–6 units, read by one person. The method by which separate words are read out has only been used in examinations of the effect of very slow speech on the durational patterns of words with different structures. The number of word units on this separate word list was 15.

The segmental phonemes of the test words have been composed in such a way that one can produce as many Finnish words as possible from the same phoneme grouping by doubling various vowel and consonant segments.

For example, by using the set of segmental phonemes /m/+/a/+/n/+/a/ (+/n/) a large number of two syllable words can be constructed in which the order of the different qualitative segments is the same and the effect of the different qualitative units on the durational pattern is therefore of the same kind: /mana/ 'sorcery' /maana/ 'as a land' /manna/ 'manna' /manaa/ '(he) conjures' /mannaa/ 'of manna' /manan/ 'sorcery's' /mannan/ 'manna's' /ma-naan/ 'I conjure' /mannaan/ 'into the manna'/. Correspondingly three syllable words can be constructed in which the order of the contrastive segments is the same, /manata/ 'to conjure' /manatta/ 'without sorcery' /mannatta/ 'without manna'/ etc. From this kind of series it is possible to make groups of minimal pairs for the comparison of different word structures where the phonetic frame of the contrastive segments is exactly the same. The differences in the means of sound segment durations in these word pairs may be supposed to be caused for this reason, by, in particular, the differences in word structure and not by, for example, the kind of change occurrence where in the sample words of one group there are in a certain position more phonemes possessing some known effects in the durational pattern than in the words of another group.

The total number of words in the sentence frame is 462, of which 79 are 'nonsense' words and 383 'real' words in meaningful sentences. The number of measured words is thus about 4500. The nonsense-word frame sentence was in all cases »Tätäkö _ tarkoittaa?» »Does _ (the word) mean this?» This sentence arrangement was chosen according to its stress pattern. In the frame sentence of the type: »Mitä _ tarkoittaa?» 'What does _ (the word) mean?' the natural stress of the sentence falls strongly on the word under consideration and partly modifies the durations of its phones. On the other hand, the structure of the sentence »Tätäkö _ tarkoittaa?» presupposes that the test word is already familiar and is therefore pronounced with weaker logical sentence stress. In traditional terminology the test word is not in the sentence as a 'psychological predicate' which normally gets the strongest subjective sentence stress (Sadeniemi 1949, p. 91–95).

The sentence frames of words really in existence are not identical, but short meaningful sentences have been used as frame sentences. In these the writer has, as far as possible, tried to place the word under consideration in similar, neutral stress positions in each sentence. The sentences are not all grammatically complete (with both subject and predicate) but the group includes meaningful and independent sentence parts which may be used as independent expressions in colloquial speech, such as, »Tuon *takan* kohdalla.» 'Close to that fireplace'. »Kaikki *pyssyt* pois.» 'All guns away!' »Hyvin *laakaa* kukkulaa.» 'Of a very large hill.' All the sentences are constructed in such a way that the word under consideration comes within the sentence and represents, on its own, one typical unit of Finnish speech, a speech measure (Finnish *puhetahti;* Sovijärvi 1946). The logical content of the sentence is such that the first and last measure are strongly stressed and the measure formed by the test word is tertiary stressed. Since the test word forms a speech measure it also contains a weak sentence stress (cf. chapter 3. 12. 1.). Thus the choice of the middle (sometimes the second in a four measure sentence) and least prominent speech measure as the sentence position for the words selected for the analysis does not mean that words in this position are completely unstressed and phonetically reduced. All the test words are phonetically »complete» and comprehensive even if separated from their environment, i.e. one can say that the phonetical code is full in the chosen position. The words constituting the final primary-stressed speech measure of the sentence are, instead, often strongly reduced: the intensity of the voice is noticeably weakened after the first syllable of the last measure and the final part is often (almost regularly with informant number 4) a completely voiceless whisper.[1]

It could be claimed that the only real material for examining, for example, quantities, would be natural spontaneous speech. However, this includes a great number of factors influencing duration which are apt to cover up the durational patterns characterizing the structure of the full phonetic code which it is the job of the analysis to reveal. Examination of the amount of reduction allowed in continuous speech and in a natural speech situation by the redundancy on various levels is fundamentally a problem of a different character.

3.3. *Technical procedures*

The material intended for analysis was recorded in the small studio of the Departement of Phonetics at Jyväskylä University. The studio does not elimi-

[1]The voicelessness of the 'tail' of an utterance is not a consequence of an extraordinarily careless speech habit but a very common feature in spoken Finnish.

nate echo totally but the reverberation of sound is quite effectively reduced. An Altec 29 A condenser microphone and a Revox G 36 tape recorder were used for the recording. The recording speed was 19 cm per second. From the studio the reader could during the recording stop the tape himself by means of a remote control switch when necessary (e.g. when coughing or when getting muddled in his reading).

All test sentences were analyzed by an ink-writer to which Frøkjær-Jensen's Trans-Pitchmeter and Intensity meter had been attached. The four-channel Elema-Schönander Model 34 Mingograph was used for the ink-recording. The fundamental frequency curve, the 'duplex oscillogram', the full frequency intensity curve and the intensity curve with high-pass filtering of 500 cps were recorded synchronically with this equipment. Some of the material was also analyzed with a Kay Electric Co. Sound Spectrograph Model R.

So far as the definition of the limits of phonetic sound segments is concerned, the most important curves are the duplex oscillogram and both intensity curves. The duplex oscillogram is not the original oscillation of the sound and does not describe the original wave-form, but it does separate from each other very effectively speech sounds with energy maxima situated on different frequencies. This is achieved by the whole signal being half-wave rectified; the DC-voltage is mixed with this as a negative part, obtained by rectifying the HP-filtered (cut-off frequency about 1500 cps) component from the original oscillation. By suitable adjustment of the voltages of these two channels, those sound segments with characteristic energy maxima on very low frequencies (nasals, *d*, *v*) show in the duplex oscillogram as only positive peaks. The more energy there is in a sound on higher frequencies, the more the oscillation of the whole duplex oscillogram bends downwards from the zero line. Sound segments containing a considerable amount of energy only in the high frequency ranges are seen in the duplex oscillogram as downward dips of the line (see the *s*-segments in the mingograms on p. 54–55). Especially in the mingographic registration the duplex oscillogram gives more information than a normal oscillogram, because the galvanometres of the mingograph cannot alone register frequencies higher than 800–900 cps.

Both mingogram intensity curves are registered by means of 10 msec integration time and a 'logarithmic' scale. In the HP-filtered intensity curve strong low frequency components of voiced sounds have been cut out so that the differences between the various voiced sounds are emphasized on the curve.

In the sample mingograms illustrated in this work, the HP-filtered intensity curve is lowest and above it is a compressed total intensity curve, the duplex oscillogram and the upper fundamental frequency curve. On the f_0- curve the variations in the tonal pattern of speech can be seen as variations of the distance between the lower edges of the saw-tooth wave and the zero-line. Exact measurement of the fundamental curve turned out to be unnecessary in this study and therefore I feel justified in not dealing with its technical background. The procedural methods for the use of the above-mentioned pieces of equipment and their functioning principles have been presented by, for example, Hadding-Koch & Petterson 1965.

The recording speed used for the mingograph was 100 mm/sec. The measurements were made with an accuracy of 0.5 mm which represents an accuracy of 5 msec. on the time scale. The variation of the tape speed of the mingograph used in this research was found, by means of control measurement, to be about 0.5 % on average.

3.3.1. »The segmentator»

For the perception tests performed with changed sound segment durations (see chapter 4) a piece of equipment was designed which makes it possible to shorten and lengthen a recorded sound segment without mechanical cutting and splicing of the magnetic tape. A block diagram of this system is shown in Figure 1. A signal, which may be a word or a whole sentence, is recorded synchronically on two sound tracks of a recording tape loop on the segmentatory equipment, so that it is repeated on each round of the tape, every 6 seconds. On the third track of the tape a pilot signal is recorded. The beginning of this is roughly placed, by ear, on the particular segment whose duration is to be changed. The moment when a recorded 1000 cps pilot signal reaches recording-playback head 3, a switch changes the output from track 2 to track 1 or vice versa. If the heads of these channels are in their original positions, there will be no change but if mobile head 1 is moved to the right, the segment which corresponds to the distance of the head from its original position is heard twice in succession: first on track 2 and then on track 1. The switch-over does not cause any audible click. If the mobile head is moved to the left, the segment corresponding to the transmission of the head gets cut off. The point at which the

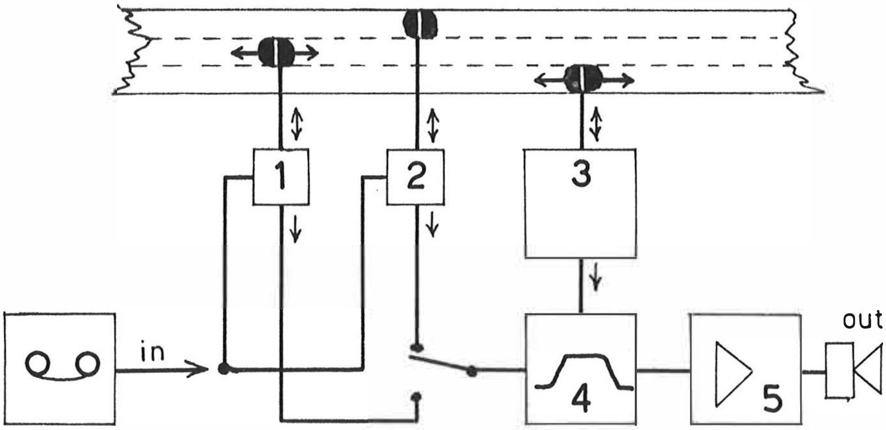


Figure 1. Block diagram of the equipment employed in changing the durations of sound segments. 1) Signal–channel with moving (record and playback) head. 2) Signal–channel with fixed head. 3) Synchronizer and switch combined with moving head on the pilot signal track. 4) Delay unit. 5) Terminal amplifier. The speech signal (a word or a whole sentence) is recorded synchronically on tracks 1 and 2. A pilot signal of 1000 cps is first auditively and then with the moving head 3 placed on the intended switching point. A signal-medial segment is then lengthened or shortened by displacing playback-head 1 to the right or to the left.

change of tracks takes place can be regulated with the help of mobile head 3 of the pilot channel. The mobile recording-playback heads are equipped with a millimetre scale. When the speed of the tape is 200 mm/sec, one millimetre is equivalent to 5 msec of time. In addition to the scales, the switching points for each test and the duration of the output are controlled by means of the mingograph. The output of the equipment is recorded on a magnetic tape in the required order and is thus presented to the listeners.[1]

When track 1 is switched off, the duration of word-initial and word-final sounds can be changed by transference of the pilot head 3 alone. But in that case, the beginning of the cut sound oscillation, and correspondingly the end, should be delayed, in order to avoid the impression of an initial or final plosive. Such an impression could easily be formed in auditory observations owing to a sudden increase or breaking up of sound energy. The most useful delay time turned out to be 50 msec O-max. The intensity curve of a sound delayed by this standard is very similar to the final abatement of the intensity of the original vowel: the curve, on descending, slopes gently at first, very steeply for 2–3 centiseconds

[1] Analogous electronic tape cutting and splicing methods are described by Huggins 1968 p. 4–5 and Pohlink 1968.

and finally slopes gently again. Also, the shortened and delayed end of a vowel cannot be separated auditively from the original one. The situation is slightly different at the beginning of a word because normally at the onset of word-initial vowels there may be a group of irregular slow glottal flaps which cannot be simulated with the methods used here. However, the delayed beginning was found to be auditively more natural than an undelayed one.

3.4. *Definition of phonetic segments*

The analysis and modification of the continuum of sound waves into an optical form in which acoustic patterns are visible and recognizable is no longer nowadays one of the central problems in the study of sound quantities. The difficulty in analysis is rather in the definition of the borders of speech-sounds to be measured. In traditional kymographic measurements the aim was to find out, by the kymograph's recordings, those phenomena which were caused by certain familiar articulations (the opening and closing of the nasal path, the beginning and end of vibrations in the larynx etc.). Kymographic duration measurement was, in fact, duration measurement of certain familiar articulations. In the case of duration measurements by means of registration of the acoustic output, the procedure for segmentation may, however, be subjected to arbitrary interpretations. On the one hand, real articulatory facts and on the other hand, perception and phonemic segmentation of the speech code may be used as a frame of reference. Generally the boundaries of phonetic segments are fixed on the maximal points of change in the diagram representing the acoustic signal; these are known to represent major points of transition in the activities of the vocal mechanism. These points of transition are supposed to represent, to some extent, the boundaries of the segments of acoustic signals which, on the level or perception, contain the fundamental characteristics of a speech sound. However, the speech signal is not a sequence of separate units with distinct boundaries. Neither is it a series of mere fluctuating and overlapping patterns.[1] In the acoustic signals, as well as the articulation level, one can observe quite stable stages with intervening transitions of greater or shorter length. The longer or shorter transitional parts between 'target points' of segmental

[1] On the relations between ideal phoneme sequences, sound segments and sound feature segments, see Fant 1967, p. 97–98.

phones have on the level of perception a double reference, because they include cues for the identification of both the previous and following phoneme. This means that segmentation, if it is taken as dividing the signal up into successive fragments, is impossible on the basis of the perceptual cues. If the overlapping perceptual segments are measured the sum of segment durations would be considerably greater than the total manifested duration of the utterance formed by the segments (Elert 1965, p. 82). The segmentation methods described and applied by Peterson & Lehiste (1960) and Elert (1965) and adhered to in the present study amount principally to a definition of articulatory segments based on the information given by the acoustic signal. Although these segments do not correspond to 'perceptual phones', i.e. segments which contain the identification cues for a corresponding segmental phoneme (see Elert 1965, p. 82), they still obviously have a certain kind of subjective or perceptual reference: the isolated segments can, at least, in most cases, be taken as those segments of speech signals which, in subjective evaluation, correspond most closely to the moments of producing phonemic speech units. Details of segmentation, such as the placing the burst of a plosive and the following noise with the duration of the plosive or with the duration of the following vowel, are only of academic interest in this sense. The purpose of this sort of 'microscopic' definition of segment boundaries is not to fix any functional boundaries at certain points on curves, but to find some fixed points which appear as regularly as possible to be used systematically in phonetic measurements.

3.4.1. Descpription of segment boundaries

Since most measurements have been made on mingograms (with the exception of sequences of vowel + *j* + vowel which have been measured on spectrograms) only those features which show the boundaries of sound segments in the curves produced by this equipment will be treated in detail in the following study: in the duplex oscillogram, in the full frequency intensity curve and in the HP-filtered intensity curve. A sawtooth curve on the mingogram indicating variations in the fundamental frequency has in no case been taken as a definite indication of a segment boundary.

3.4.1.1. Plosives

In a kymographic registration with a mouth-funnel, it is possible to obtain information about the closure articulation of the plosive at the so-called absolute beginning with the help of the change of the static atmospheric pressure in the mouth channel which is caused by the movement of the organ of articulation to a closure position. In an acoustic analysis the corresponding duration of articulation of a plosive in this kind of position cannot be found since implosive articulation does not cause any audible or instrumentally discoverable crack. The only measurable duration of the plosives in this position is the interval between the explosive burst of a plosive and the beginning of the fundamental oscillation which interval is very short in Finnish owing to the lack of aspiration. In very slow speech the interval between the beginning of an explosive peak and the first fundamental period may, however, be 30–50 msec; the burst and the subsequent voiceless interval is slightly longer with a velar-plosive *k* than with the plosives *t* and *p*.

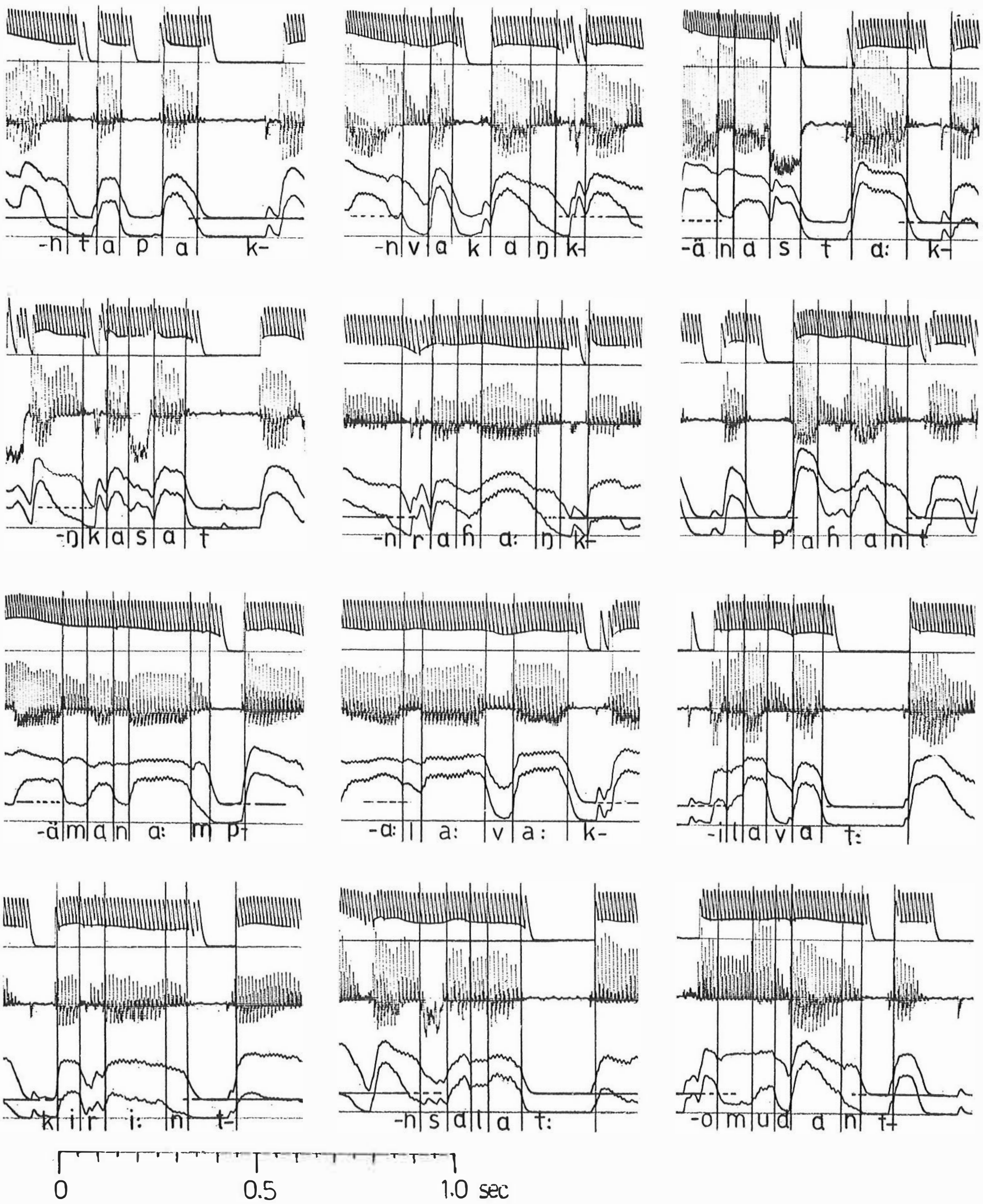
In cases where plosives appear at the absolute beginning (individual words read slowly and words in the perception tests) the common duration of an explosive burst and of the very short voiceless interval which may follow it have been taken as a duration of the plosive on the tables and graphic diagrams. In cases where a plosive appears between voiced sounds, the beginning of the plosive has been determined as the point where the intensity curves descend steeply and the amplitude of oscillation quickly abates indicating the placing of the organ of articulation in the closure-position. This corresponds to the point, on a spectrogram, when the energy suddenly grows weaker on the formant area and the last lines indicating formant transitions disappear (compare Fischer-Jørgensen 1964, p. 182; Öhman 1966, p. 984.). The definition of the boundary between a plosive segment and the next vowel has over the years been the subject of different interpretations (see Elert 1965, p. 80; Wiik 1964, p. 35), and there appears to be no common practice in the definition of boundaries even in recent measurements based on acoustic material. On some occasions the explosive burst of the plosive has been taken as the boundary between the plosive and the next vowel, while on other occasions the boundary has been fixed at the appearence of the fundamental oscillation or the formant structure of the vowel. The aspiration or voiceless period following the noise burst of the plosive have been counted as belonging to the duration of the vowel by, for instance, Wiik (1965, p. 35) and Wiik & Lehiste (1968, p. 569). Obviously the boundary of the segments has been determined as the middle of the 'releasing spike' of the plosive also in a couple of studies in which the definition of boundaries has not been described in detail but in which the criteria established by Peterson and Lehiste (1960) are said to be followed. However, the abovementioned writers have used two separate methods of measurement in their study 'The Duration of Syllable Nuclei in English': the syllable nucleus is thought to begin 1) at the centre of the releasing spike and 2) with the onset of voicing immediately after aspiration. The duration of the vowel measured according to the first alternative is, for example, in the English word *tuck* as long as in the word *dug*, but if the duration of aspiration is not included in the vowel, the duration of the syllable nuclei of the words *tuck* and *duck* is approximately the same. The first alternative would also mean that a voiceless aspirated initial consonant would have a very strong lengthening effect on the following vowel; the latter alternative would mean a shortening effect corresponding to general observations on the effect of plosives(see Peterson & Lehiste 1960, p. 701). The boundary between the plosive and the following vowel has been set after the aspiration of the plosive or after the voiceless segment following the burst: by House (1961) in his study on English quantities, by Fischer-Jørgensen (1964, p. 182) in her research on Danish and by Elert (1965, p. 83) in his research on Swedish quantities, and also by Gårding (1967, p. 94) in her research on Swedish juncture. Öhman (1966, p. 984) has considered the first transient spike of the release as the CV boundary of a voiceless plosive, but considered the voice pulse in which the higher formants reappeared in the spectrograms after the closure segment as the boundary of a voiced plosive.

In the present study the point at which a fundamental voice oscillation commences after the explosive burst of a plosive and the intensity curve rises steeply has been taken as the CV-boundary of the plosives. In the case of the only normally voiced plosive /d/ the corresponding CV boundary is where amplitude of the sound oscillation suddenly increases and the intensity curve rises steeply. In the spectrogram this corresponds to the point at which the vowel formants reappear. Because the Finnish plosives are normally unaspirated, the choice of the CV boundary for the plosives has no great influence on the results of measurements. Reappearance of vocal cord oscillation and the formant pattern of the vowel have been chosen as the boundary here because the same criteria can also be used in defining the boundaries of several other consonants.

3.4.1.2. Fricatives /s/ and /h/

The Finnish apicoalveolar grooved fricative /s/ is not always voiceless, particularly when its duration is short. Between voiced sounds it is quite often partly or wholly voiced. Because the termination and beginning of vocal cord vibration cannot be used as boundary criteria, the only useful criterion is the existence of noise. The appearance of high frequency fricative noise shows in the duplex oscillogram as a deviation of the baseline from the zero-level synchronically with a clear decrease of the amplitude of oscillation. In the HP-intensity curve at the beginning of the fricative segment there is a quick downup motion which can be considered, together with the duplex oscillogram, as the sign of the initial boundary of an *s*. The same kind of hollow appears at the end of the fricative segment before the beginning of the oscillation of the following vowel. This steep rise intensity together with the sudden increase in amplitude and the change in the wave-form of the duplex-oscillogram have been taken as the CV-boundary of *s* (cf. Peterson & Lehiste 1960, p. 694).

Defining the boundaries of the allophones of the fricative /h/ is more difficult because it has no typical noise frequency area due to the fact that its noise substance mostly appears in the area of the higher formants of the adjacent vowels. Peterson and Lehiste's (1960, p. 696) boundary criterion cannot be used in mingogram measurements: i. e. the termination and beginning of the voiced oscillation in the F_1-area. Nor did this seem to be an adequate criterion in the spectrograms with a group of *h*-substance words used experimentally to establish criteria. The best characteristic of *h* turned out to be a decrease in intensity at the beginning of the glottal constriction related to a clear reduction of amplitude in the duplex oscillogram (cf. Fischer-Jørgensen 1964, p. 182). Correspondingly, the CV-boundary of *h* is set at the point where the intensity curve rises rather steeply and the amplitude of oscillation begins to increase. In some cases, the *h*-segment is also related to a slight descent of the fundamental frequency, but this feature is not systematically repeated and cannot be considered a sufficiently accurate boundary criterion.



3.4.1.3 Nasals /m/, /n/, /ŋ/

There are no difficulties connected with the defining of boundaries of nasals in word-initial and word-medial positions (cf. Peterson & Lehiste 1960, pp. 695–696; Elert 1965, p. 85). The initial and terminal points of nasals are seen

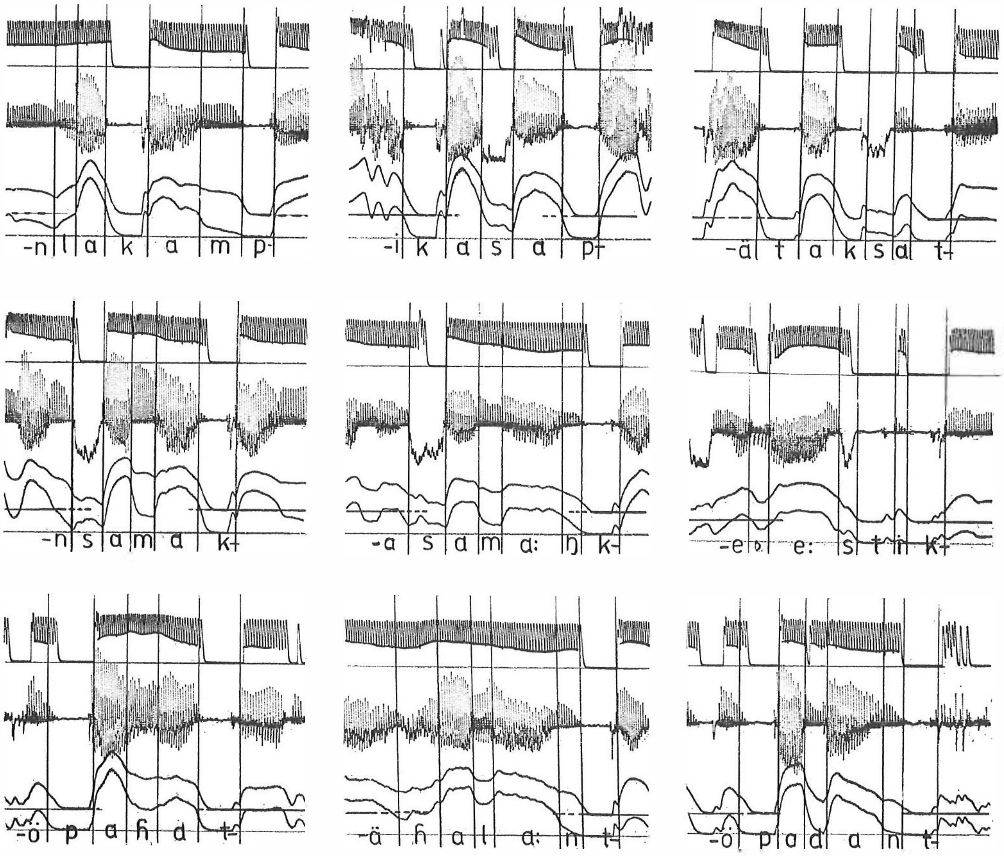


Figure 2. Mingograms of the test words spoken by male speaker 5 and by female speaker 9. At the top of the fragments of the mingogram band there are fundamental frequency curves, below duplex oscillograms and undermost the two intensity curves: the upper curve is a full frequency curve and the lower a high-pass filtered intensity curve. Both of the intensity curves are registered in a compressed scale with an integration time of 10 msec. The recording speed of the ink-writer was 100 mm/sec. All of the sample words are spoken in a sentence-medial position. In the phonetic representation beneath the intensity curves, letters with a line before or after (e.g. -n, n-) show the last resp. first sounds of surrounding words.

clearly in the duplex oscillogram as a change in the wave-form and as a sudden descent and rise in the HP-intensity curve. However, the boundary of word-final nasals is not always clear: the vowel is often nasalised in its latter part and the gentle descent of intensity often starts at the very beginning of the vowel segment. This obviously means that the activity of the velum in such a non-stressed position is reduced; the velum may already be lowered a little at the beginning of the vowel-nasal sequence and its motion is carried through more slowly than usual to the end of the sequence. However, a steeper descent of the HP-intensity curve can be seen at the end of the sequence which may show a more sudden darkening of the quality caused by the closing of the mouth path. There is often a simultaneous movement of the oscillation in the duplex oscillogram to the low frequency side. Descent of the HP intensity curve and a simultaneous change in the waveform has been taken as the initial boundary of word-final nasals. The end of word-final nasals is not clear either, particularly before a plosive where an abating oscillation may continue until the plosive burst. In these cases, the boundary between nasal and plosive occlusion is, however, generally visible as a clear descent of the total intensity curve.

3.4.1.4. Semi-vowels /v/ and /j/

The boundaries of *v* are particularly clear on the mingograms: the clearest characteristic of the initial boundary is a steep descent of the intensity curve and also a related decrease in the amplitude of the duplex oscillogram. Correspondingly, the final boundary of *v* shows just as clear opposite changes in the intensity curves and in the duplex oscillogram (cf. Peterson & Lehiste 1960, p. 697; Elert 1965, p. 85). On the other hand, definition of the boundaries of the palatal semi-vowel *j* is more difficult. In most cases the curves of the mingograms did not give any information about the boundaries between *j* and vowels surrounding it. For this reason, all measurements of *j* have been made by means of wideband spectrograms. But even so, it is not easy, without ambiguity, to find the boundaries of *j* on the spectrograms because the acoustic formant transitions may slope extremely gently. There are cases in which it is possible to observe, in a short j-segment, a brief steady state where F1 and F2 have reached their respective minimum and maximum values and the formant energy has clearly weakened. But in many cases inter-vocal *j* can be seen on the spectrograms only as a continuous formant glide without any clear steady states. The boundary used by Öhman (1966, p. 984) has been selected as the segment boundary in this study: i.e. the point at which the transitions seemed to be maximally rapid. This in itself arbitrary boundary criterion differs from those used with other consonants, but it seemed to be the only practical one under the circumstances. The effects of various alternative principles of boundary defining on the results of measurements of *j* is dealt with on page.

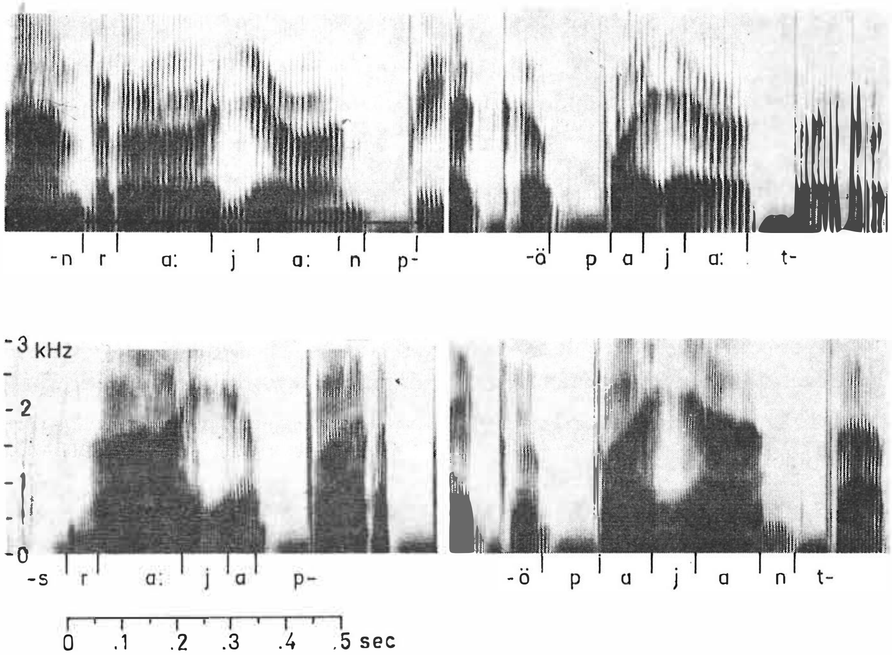


Figure 3. Samples of sound spectrograms illustrating the segmentation procedure used when defining the boundaries of the voiced semi-vowel /j/. The upper two spectrograms are made of the sentence medial words *rajaan* and *pajaa* spoken by male speaker 5, the lower two are of the respective words *raaja* and *pajan* spoken by female speaker 9.

3.4.1.5. Liquids /l/ and /r/

The boundary of a lateral does not cause any difficulties: the beginning and end of *l* shows clearly on the mingogram as a change in the wave-form of the duplex oscillogram and, in particular, as a steep jump in the HP-intensity curve. The most usual allophones of the Finnish /r/ phoneme are rolled and flapped frictionless apicals and a slightly fricative apical *r* in which the trill with the tip of the tongue can be negligible or completely absent. The beginning and the end of the flapped and fricative *r* are easy to define by steep descents and rises of the intensity curves. On the other hand, there may occasionally be problems in defining the end of a rolled *r:* the sound does not end with the last clear stroke of the tip of the tongue, but is followed by a kind of incomplete stroke which shows in the intensity curve as some kind of gently sloping step before a vowel.

Samples of the methods used in this study for the definition of the boundaries of various types of segments are shown in the spectrograms and mingograms on pages 54, 55 and 57. The samples are taken at random out of the speech of one male (speaker 5) and one female informant (speaker 9). The bits of mingographic tracings of speech include illustration of segment boundaries used in this work for all consonant segments exept for /j/ the boundaries of which are defined only spectrographically. The definition of segment boundaries and the resulting duration values also depend, to some extent, on the technical adjustment of the research equipment. When the mark level of the sound spectrograph or the recording amplitudes of the mingograph galvanometres have been adjusted to a greater value, the duration of voiced segments measured tends to be longer than on tracings made with smaller mark level and recording amplitude.[1]

3.5. *Statistical description and analysis*

The artihmetic means ($\bar{X}$) and standard deviations (s) are calcuted from the durations of measured sound segments of each word read by the 10 speakers.[2] In calculations of the average durations of sound segments of word structures where several words are included in the set representing the structure, means and deviations have not, however, been calculated from the means of words but from the original individual results of measurements. The significance of the difference ($\bar{d}$) in the means of durations of sound segments in different structures, different sound categories or in other groups has been measured by the so-called t-test, but the t-test has only been performed within groups which have been considered linguistically comparable. The durations of various sound segments in fully identical phonetical frames, durations of sound segments in minimally differing word structures with identical sound composition etc. are accepted as comparable. The difference of the means has been considered as strongly significant, if the risk level of probability of difference (p) shown by the t-test is less than 1 % ($p < 0.01$), as significant if

[1] Owing to the characteristics of the analysis equipment on the one side and of the sound-waves of speech themselves on the other side, the 'absolute' accuracy of an individual measurement is hardly more than $\pm$ 10 msec, even though the actual measurement is performed with an accuracy of 0.5 mm or 5 msec (cf Lehiste 1970 p. 100).

[2] The values of the mean durations and standard deviations in individual durations of sound segments in the individual words are to be found in the unpublished appendix in the library of Jyväskylä University.

$p<2\ \%$ and almost significant if $p<5\ \%$. The significance of the difference in the comparison tables of means is shown either as a percentage or by an asterisk so that one asterisk (*) = almost significant ($p<5\ \%$), two asterisks (**) =significant ($p<2\ \%$), three and four asterisks strongly significant (*** = $p<1\ \%$; **** = $p<0.5\ \%$). In some comparisons even the 10 % and 20 % levels of significance which are not significant in the statistical sense, but which can be considered indicative in a particular connection, were marked with a percent figure. In the comparisons between different speakers and sexes a one-way variance analysis was used instead of the t-test.[1]

Representatives of the so-called phonometric school have stressed the use of statistical and particularly of variation-statistical methods in examining sound quantities. According to the phonometric method experienced researchers shoud first classify the sounds of the test for analysis subjectively according to selected quantity categories, for example, short, long and overlong. »Erst nach dieser Bestimmung nach Längen und Kürzen haben Messungen einen sprachwissenschaftlichen Sinn, indem sie zeigen welche Variationen absoluter Lautdauer in der betreffenden Sprache für die Längen und welche für die Kürzen beobachtet werden und wie diese absoluten Werte sich um Mittelwerte scharen – –» (Zwirner & Zwirner 1937, p. 99).[2] Individual measurement results would not, according to the phonometric school, have any significance either, because on this basis one cannot decide whether a language community interpets a sound as, for example, long or short.

However, what the listener »hears» when he receives the acoustic signal is a particular sound and contrast between the durations of consecutive sound segments, not the means and distributions of durations in separate positions in an arbitrary text. Those cues which facilitate the decoding of an acoustic speech signal for both a trained phonetician and a naive listener, must themselves appear in the acoustic signal and not in some non-existent means. The

[1] A detail explanation of statrstical description and analysis methods and their mathematical patterns here is given by McNemar (1963), p. 20– (standard deviation), p. 98– (the t-technique) and p. 252– (analysis of variance). The p-values of this work are defined according tp Henrysson's (1965, p. 124) t-distribution table. For the evaluation and representation of the significance of differences in means see also Elert 1965, p. 87. The calculations fos statistical description and analysis were made in the Computer Centre of Jyväskylä University under the guidance of Mr. Keijo Lipén.

[2] See also Maack 1949, 1931 and Magdics 1961, especially p. 27, for a deseription of the phonometric method.

calculation of duration means of sound segments is valid only if the units to be compared are in similar frames affecting the timing of segments: in the same stress position, in the same sentence intonation and tempo position.[1] The invariance which the phonetician endeavours to reveal by mathematical methods can even then only partly imply the absolute character of measured phenomena (in the duration of a segment in this case) – it can also partly imply the contrastive relations between the successive units in the flow of speech.

3.6. *Intrinsic duration characteristics of segmental phoneme units*

Factors on several different levels affect the objective measurable duration of the sound segments of some produced concrete expression. It is supposed in the present research that the effect of speech speed or tempo has, in each group of examined words, been on average the same in different frame sentences. By selection of speakers and by use of the mean of the sound durations of expressions read by ten separate persons in different comparisons, an attempt was made to keep individual pronunciation habits, incidental factors and geographical and social factors which might affect the speech utterance uniform. Of factors more closely related to language and depending on the particular structure of each language which affects the durations of sounds, an attempt was made to level out the effects of stress and intonation (or the effects of rhythmical phrasing of an utterance) by using the same kind of sentence frame. There remain two kinds of factors affecting the duration of phonemes; of these one kind could be called paradigmatic and the other syntagmatic: paradigmatic factors affecting the duration of phonemes mean, in this study, those factors which are caused by particular characteristics of each segmental phoneme. These do not concern, in their effects, only the sound segment in question, but also include the effect of a certain phone on the duration of the sound before or after it. Syntagmatic factors are those which depend on the

[1] According to the phonometric method the distibution of the durations of, for example, German single and double vowels is described in the form of two Gaussian curves, between which there is considerable overlapping. It has been shown, however, that in the same sentence-phonetic position German long (tense) vowels are always longer than the corresponding short vowels (Delattre & Hohenberg 1968 p. 380), in other words, overlapping does not occur at all.

structure of a larger phonological unit (e.g. syllable, measure or word) regardless of which vowel and consonant phonemes of the paradigm appear in different positions within the structure. For example, the nasal /n/ differs, in regard to phonetical duration, from the other consonants in certain proportions and it has, when compared to these, a known effect on the duration of the preceding and following segments in any structure. These kinds of durational differences and durational proportions are, therefore, paradigmatic in character. On the other hand, a word whose structure is, for example, single vowel + single consonant + single vowel differs in its durational distribution from other structures, regardless of what qualitative sound units fill the various positions in the structure. This kind of durational factor may be called syntagmatic.

It cannot be categorically said of either a paradigmatic or syntagmatic durational factor that it has a physiological cause or that it is linguistic and learned. The inherent durational characteristics of segmental phonemes which are valid (mutatis mutandis) in any structure where the phoneme in question appears, can be affected by the complexity of articulation movements accompanying them, various action speeds of different articulations, the difference in motor channels etc. But the different inherent durational effects of a phoneme can equally well reflect timing distinctions which are linguistically conditioned, intended and learned. Correspondingly, syntagmatic durational factors are not necessarily functions of merely intentional and learned timing patterns, but these are also affected by certain restrictive physiological or motor factors influencing the production of speech units larger than a phoneme.

The purpose of this chapter is to show to what extent the objective duration of sounds depends on the nature of the segmental units in the structure. The main purpose is simply the description of the prevailing durational differences. The reason for these cannot be found by means of the methods used here. First each qualitative segment's own relative and absolute duration is examined and this is followed by an examination of the absolute and relative effect of a qualitative segment on the durations of adjacent segments.

3.6.1. Vowel durations

In analyses of various languages it has been shown that in identical phonetic environments certain vowels are, on average, longer in their durations than others. In the Finno-Ugric languages which have been studied, just as in the many well-known Indo-European languages (see Elert 1965, p. 122), the duration of vowels has generally been shown to depend on, for example, their degree of openness, so that more open vowels have been shown to be longer in duration than less open ones. In their study on Hungarian, Meyer and Gombocz arrange the short vowels in the following order, according to duration: /i u ü o e ö ä å/ and the long vowels, similarly: /ī ū ē ǖ ō ȫ ā/ (Meyer & Gombocz 1909, p. 133).[1] The average arrangement of Lappish vowel duration, according to Äimä (1914, p. 147), is /i u̹ u o ɛ̜ a̹ e ȧ ạ ä/.[1] This order has been obtained by grouping the vowels in each of the four quantity classes according to duration and by adding the rank numbers obtained for each vowel. Liiv has grouped the vowels of stressed syllables into three groups: the shortest are the vowels /ü i u/, the following group is formed by /e ö e̯/ and the longest in duration are /ä a o/ (Liiv 1961, p. 417).[1] If however, in accordance with Äimä, one groups the durations of the individual vowels in different quantity degrees given by Liiv according to the sums of their rank numbers, the order is /ü i e̯ u ä e a o/. The corresponding order of the mean vowel durations of the three quantity degrees in Estonian, calculated according to Lehiste's (1960 Table on p. 52) published means, is /i ü e̯ ö e a o ä u/.

In the Salmi dialect of Karelian (spoken in the parish of Salmi on the North-east side of Lake Ladoga) the durational order of the short vowels in 1st, 2nd and 3rd syllables have been calculated from Donner's (1912, p. 8) measurement results: /i u e ä y a o/ (in Donner's material the vowel /ö/ only appears in the first syllable in which it is the second longest after /a/ in duration). The arrangement of the vowel duration of the South Ostro-Bothnian dialect of Finnish according to first-syllable short vowels is, starting with the shortest /e i u o y a ä ö/ (Laurosela 1922, p. 158). In his study on Finnish and English vowels (1965, p. 115), Wiik has calculated the average durations of single and double monophthongs in the first syllable of disyllabic words in three acoustic vowel types: /low F_1/ vocoids (articulatorily high vowels) /i y u/, /mid F_1/ vocoids (mid vowels) /e ö o/ and /high F_1/ vocoids (open vowels) /æ a/. The average vowel durations of the groups are as follows:

[1] The phonetic representation of the vowels is in accordance with the original text.

	single, cs	double, cs
/i y u/	9.5	23.1
/e ö o/	10.0	23.5
/æ a/	10.3	22.8

The durational order of individual vowels is, in accordance with the values given by Wiik, in the single vowel group /i e y ö u a o æ/ and in the double vowels group /e o i æ u a y/.

To throw light on the durational proportions of vowels, all the vowels of Finnish were read in an identical sentence frame in two nonsence word structures, pVVpV and pVppVV, which were composed so that in each structure the first and second syllable had the same vowel, for example *paapa*, *piipi*, *pappaa*, *pippii* and so on. The frame sentence was *Tätäkö* _______ *tarkoittaa?* 'Does _____ mean this?' Thus, before and after the first syllable vowel there was always the labial plosive *p*, before the vowel of the second syllable *p* and after it the dental plosive *t*. The measurements of all sound segments of nonsense words are shown in table 2. When the vowels of each group in the tables are arranged according to their average duration and the rank numbers of each vowel obtained in the different groups are added together, the vowels are arranged in the following durational order: /ä a o u ö y e i/. Therefore the longest are the open illabial vowels /ä/ and /a/, the second group is formed by the labial back vowels /o/ and /u/ and the labial front vowels /ö/ and /y/ in such a way that in each pair the more open vowel is longer in its duration. The shortest in this rank order are the illabial front vowels /i/ and /e/.

The relative, intrinsic duration values of vowels (see Maack 1949; Elert 1965, p. 129–), or the ratios of the durational means of different vowels to the mean of all vowels in separate structures, are calculated in Table 3.

If the means of the intrinsic durations in table 3 are examined, the durational order of vowels is seen to be about the same as the rank-order already presented. The open vowels /ä/ and /a/ are the longest, followed by the labial vowels /u o y/ and /ö/ and the illabials /e/ and /i/. The durational order of sounds shows that, at least in connection with labial consonants, the lip-rounding articulation clearly has a lengthening effect on vowel duration. The most obvious reason is that during bilabial closure the lips cannot settle in the protruded rounded position during the consonant as happens when a labial

TABLE 2. Mean durations ($\overline{X}$) in milliseconds and standard deviations (s) of different sound segments in nonsense-words /p V V p V/ and /p V p p V V/ with various vowels in the V-positions.

Nonsense word frame /p V V p V/		p–	–vv–	–p–	–v	Nonsense word frame /p V p p V V/		p–	–v–	–pp–	–vv
/a/	$\overline{X}$	113	180	83	48	/a/	$\overline{X}$	92	79	227	142
	s	16	20	13	8		s	16	10	18	18
/ä/	$\overline{X}$	117	180	82	48	/ä/	$\overline{X}$	94	84	222	144
	s	31	22	17	10		s	12	11	20	20
/o/	$\overline{X}$	124	168	98	44	/o/	$\overline{X}$	100	79	239	134
	s	16	19	6	12		s	10	12	32	17
/ö/	$\overline{X}$	120	168	90	41	/ö/	$\overline{X}$	94	81	227	127
	s	10	16	12	11		s	11	11	23	16
/e/	$\overline{X}$	111	163	88	41	/e/	$\overline{X}$	98	77	236	132
	s	20	24	12	8		s	14	12	19	17
/u/	$\overline{X}$	129	164	98	48	/u/	$\overline{X}$	119	68	244	140
	s	24	26	15	11		s	20	15	40	28
/y/	$\overline{X}$	131	168	93	48	/y/	$\overline{X}$	109	68	241	129
	s	12	30	13	16		s	16	12	35	26
/i/	$\overline{X}$	115	164	88	37	/i/	$\overline{X}$	114	74	243	138
	s	13	21	17	7		s	41	8	32	21
$\overline{X}_{all}$		120	169	90	44	$\overline{X}_{all}$		102	76	235	136
s_{all}		20	23	15	11	s_{all}		22	13	30	22

vowel follows an alveolar or velar consonant. However, Eli Fischer-Jørgensen (1964, p. 197–199) has shown in Danish that a labial plosive preceding and following a vowel has a shortening effect on both illabial and labial vowels.

The statistical significances of the durational differences of various vowel pairs are compared by means of the t-test in table 4. The differences of the means of compared vowel durations are presented in milliseconds and the levels of significance as per cent probability (p). The first columns show the differences between the vowels in the first and second syllables of *paapa*-type

TABLE 3. Relative intrinsic durations of vowels ($\overline{X}_V/\overline{X}_{a\,IIV}$) in nonsense-word frames /paapa/pääpä/ etc. and /pappaa/päppää/ etc. The absolute differences in mean durations and the statistical significance of the differences are given in table 4.

Vowel phoneme	single 1. syll.	double 2. syll.	double 1. syll.	single 2. syll.	mean
/a/	1.04	1.04	1.06	1.09	1.06
/ä/	1.11	1.06	1.06	1.09	1.08
/o/	1.04	0.99	0.99	0.99	1.00
/ö/	1.06	0.94	0.99	0.92	0.98
/e/	1.01	0.97	0.96	0.93	0.97
/u/	0.89	1.03	0.97	1.08	1.02
/y/	0.89	0.95	0.99	1.08	0.98
/i/	0.97	1.01	0.97	0.84	0.95

nonsense words; the next two columns show the differences between the short vowels of the first syllable and the long vowels of the second syllable in *pappaa*-type words. It should be noted in connection with the statistical comparison that 10 % and 20 % levels of the p-value are no longer significant, but here they can be seen as indicative and directive. The differences in the average values of absolute vowel durations are rather small, in all cases less than 20 msec, and are in very rare cases statistically significant. In the comparison of long vowels of the first syllable no durational difference of any vowel pair has been shown to be statistically significant. Only on the 20 % level can one say that /aa/ is longer than /öö/, /ee/, /uu/ and /ii/, and that /ää/ is longer than /ee/, /uu/ and /ii/. Of the single vowels in the second syllable of the structure CVVCV, /a/ is longer than /ö/ on the 20 % level, /a/>/i/ is strongly significant ($p<0.01$). /a/ is longer than /ö/, /e/ and /i/ on the 20 % level. However, the short high labial vowels of second syllable /u/ and /y/ are just as long in duration as the open vowels /a ä/ and almost significantly longer than their illabial equivalent /i/. There is, therefore, also considerable overlapping between the durations of different vowel allophones even when vowels appear in an identical frame. Accordingly the intrinsic duration of a vowel can hardly be taken as a plausible perceptual cue of a vowel's phonemic

TABLE 4. The mean differences ($\bar{d}$) in the durations of various vowels spoken in nonsense-word frames and the statistical significance of the differences as a percentage probability (p) according to the t-test.

vowels paired	nonsense-word frame /pVVpV/ double 1.syll. $\bar{d}$(msec)	p	single 2.syll. $\bar{d}$(msec)	p	nonsense-word frame /pVppVV/ single 1. syll. $\bar{d}$(msec)	p	double 2. syll. $\bar{d}$(msec)	p
/a-ä/	0	non	0	non	–5	non	–2	non
/a-o/	+12	non	+5	non	0		+8	non
/a-ö/	+12	20%	+8	20%	–2	non	+15	10%
/a-e/	+17	20%	+7	10%	+2	non	+10	non
/a-u/	+16	20%	+1	non	+11	10%	+2	non
/a-y/	+12	non	+1	non	+11	5%	+13	non
/a-i/	+16	20%	+11	1%	+5	non	+4	non
/ä-o/	+12	non	+5	non	+5	non	+10	non
/ä-ö/	+12	non	+8	20%	+3	non	+17	10%
/ä-e/	+17	20%	+7	20%	+7	non	+12	20%
/ä-u/	+16	20%	+1	non	+16	2%	+4	non
/ä-y/	+12	non	+1	non	+16	1%	+15	20%
/ä-i/	+16	20%	+16	20%	+10	5%	+6	non
/o-ö/	+1	non	+3	non	–2	non	+7	non
/o-e/	+6	non	+3	non	+2	non	+2	non
/o-u/	+5	non	+4	non	+11	10%	–6	non
/o-y/	+1	non	–4	non	+11	10%	+5	non
/o-i/	+5	non	+7	20%	+5	non	–4	non
/ö-e/	+5	non	–1	non	+4	non	–5	non
/ö-u/	+9	non	–7	20%	+13	5%	–13	non
/ö-y/	0		–7	non	+13	5%	–2	non
/ö-i/	+4	non	+4	non	+7	20%	–11	non
/e-u/	–1	non	–7	20%	+9	20%	–8	non
/e-y/	–5	non	–7	non	+9	20%	+3	non
/e-i/	–1	non	+4	non	+3	non	–6	non
/u-i/	0		+11	5%	–6	non	+2	non
/u-y/	–3	non	0		0		+11	non
/y-i/	+4	non	+11	10%	–6	non	–9	non

quality. It is also likely that the intrinsic duration of vowels does not have to be taken into account in the preparation of programmes for the automatic production or recognition of Finnish.

Even though the durational order of vowels in different languages and dialects shown by various research workers varies considerably (perhaps because the material is limited or uneven), it may be established from all the vowel series presented above that more open vowels are among the longest and closed illabial vowels are among the shortest.

Durational differences between vowels of different quality and the reasons for these differences have been examined in considerable detail with reference to, for example, English and the Scandinavian Languages. (English: Peterson & Lehiste 1960, House 1961 etc.; Danish: Eli Fischer-Jørgensen 1955, 1964; Swedish: Elert 1965, Lindblom 1967). House has shown in his study based on bisyllabic nonsense utterances that in groups of both tense vowels and lax vowels the average duration of vowels increases from the more close to the more open vowels (p. 1176). The dependence of vowel duration on the degree of openness of vowels has also been shown by Eli Fischer-Jørgensen, in her work on Danish, by means of nonsense words. According to this, the differences in average duration of the groups made in accordance with the degree of openness of vowels /i y u/, /e ø o/, /ε œ ɔ/ and /a/ are also statistically significant. The number of studies in which mutual dependence between vowel duration and degree of openness has not been established, is very small (cf Elert 1965, p. 122). However, Such results have also been obtained among studies which have used modern methods of acoustic analysis. Thus Fintoft asserts in his study on Norwegian sound durations (1966, p. 35): 'There does not seem to be any basic difference between the duration of the vowels *a*, *i* and *u*. Only in certain few words is the duration of initial and final *a* greater than the duration of *i* and *u* in analogous compounds. The average duration of the vowels is the same.' The relationship between the degree of openness and durational differences between vowels also seems to be irregular in Persian (Strain 1969).

It may be said of the grouping by duration of Swedish vowels performed by Elert (1965, p. 122: /i y o/, /e u å /, /ä ö a/) and of the durational order of Finnish vowels in this study that they are roughly equivalent to the vowel series extending from more close to more open. On the other hand, the relationship between degree of openness and vowel duration in the series of Finnish double vowels in Wiik's study is almost the contrary: the average duration is longest in the group of mid-vowels, slightly shorter in the group of close vowels and shortest in the group of open vowels. This unexpected relationship could be compared to the factor indicated in the present study, that the durational differences between double vowels are, on average, statistically less significant than the durational differences between phonetically short single vowels. The formulation of the phonetical duration of Finnish double vowels is clearly affected by factors seeking to level out the durational differences caused by various degrees of vowel openness.

The dependence of vowel duration on the degree of openness of the vowel was first demonstrated by E.A. Meyer who supposed that the shortening of more closed vowels was a compensation for the energy needed for greater tongue height (Meyer 1903, p. 39 f). Eli Fischer-Jørgensen supports Meyer's physiological explanation, but in reverse: in reality, according to Fischer-Jørgensen, the open vowels are farther from the rest position and thus require greater effort (anstrengelse) than the close ones (Fischer-Jørgensen 1955, p. 54). If the duration of a vowel were to be lengthened by the articulation being drawn away from the rest position, one would expect the mid vowels and, with regard to frontness, the central vowels to be shorter in duration, but this does not, however, hold true. In her later study, Fischer-Jørgensen (1964, p. 207) shows that the duration of a vowel depends on the extent of the movement of the speech organs required in order to come from the vowel position to the position of the following consonant. The effect of the preceding consonant is less clear, according to Fischer-Jørgensen, but seems to go in the same direction. In her view, the motor command for the timing may be the same irrespective of the quality of the vowel, but the execution of the command may be delayed owing to the movements to be made. B. Lindblom considers that the primary reason why open vowels tend to be longer than close vowels is the dynamic behaviour of the mandible: the extent of mandibular movement is greater for open vowels even during the occlusion of the initial consonant but in the transition from the vowel to the final consonant the mandibular off-glide movement progresses so slowly that the contact between the articulators for the consonant is delayed in the context of the open vowel. By means of this kind of coarticulation the system manages to speed up the actualisation of adjacent sounds without having to speed up the rate of movement of the individual component gestures (Lindblom 1967, p. 21–22).

TABLE 5. Duration of diphthongs in bisyllabic nonsense words /pVxVypVy/

Diphthong	Mean	s	Diphthong	Mean	s
/ai/	199	31	/au/	185	30
/äi/	197	27	/äy/	201	34
/ei/	172	20	/eu/	179	23
/öi/	192	22	/öy/	197	35
/oi/	192	29	/ou/	178	23
/yi/	182	22	/iu/	177	26
/ui/	171	28	/ey/	196	30
			/uo/	181	40
			/ie/	162	21
			/yö/	177	25

3.6.1.1. Duration of diphthongs

The means of the durations of actual Finnish diphthongs read in the first syllable of a nonsense word are presented in table 5. The diphthongs read in the nonsense words are slightly longer in total duration than the double vowels read in the identical nonsense word frame.[1] The longest diphthongs in duration are /ai äi/ and /äy öy ey/ or, in other words, those diphthongs where the difference between the openness of components is the greatest and those where a close labial front vowel occurs as the second component. Therefore it seems likely that labialisation helps to lengthen a diphthong just as it was seen to lengthen the preceding consonant in table 2. Since only the total duration of the diphthongs is measured, the durational relationships of component parts in separate cases remain unexplained. As was expected, the longest diphthong in total duration is /äy/, where both factors causing a lengthening of duration appear: the largest difference in openness and labialisation of the latter component. The differences in total duration between the various diphthongs are fairly small: the durational difference between the longest and the shortest is only 39 msec (i.e. /ie/ is 19 % shorter than /äy/).

[1] According to Wiik (1965, p. 125), the duration of diphthongs corresponds approximately to the duration of double vowels. The duration of the components of diphthongs would corresppndingly be about the same as the duration of the equivalent single vowels. A detailed analysis of the components of diphthongs and of vowel sequences which include a syllable boundary as well as the trasition from one to the other has been performed by Karlsson (1970).

3.6.2. Duration of consonants

The average durations of different consonants in nonsense words of CVCV, CVCVC, CVCVV, CVCCV types in intervocal position are given in table 6 and grahically in figure 4. The ratio of the average duration of every individual consonant to the mean of the durations of all consonants appearing in the corresponding structure, in other words the intrinsic duration value of the consonant is also calculated in the table between the average durations. The means of the intrinsic duration values of single consonants appearing in the three different structures are also presented in table 6. In all the nonsence words, the initial plosive is /p/ and the vowel an open back-vowel /a/ and the final consonant of the CVCVC structure is /n/. Different consonants are inserted only in the intervocal position (e.g. *papan*, *patan*, *pakan*, *pasan*). The total means below the columns, apart from the fourth column, are the mea nsof all the measured cases and not the means of themeans.

The significance of the durational differences between consonants is tested by comparing the means of different consonants in each structure in pairs by means of the t-test. The results of the test of significance are in table 7. The average durational differences between the single consonants in the various structures are to be found in the first three columns and the statistical significance of the difference is noted as a percentage probability. In the fourth column may be found the average durational differences between single consonants in milliseconds and in the right hand column are the durational differences between various geminate consonants with their significance in the only tested word structure containing a geminate consonant.

The voiceless plosives are the longest of the consonants, followed by the fricatives /s/ and /h/ and the semivowel /j/ whose boundary definitions were rather arbitrary (see chapter 3.4.). Another basis for segmentation would have shortened the measured duration of /j/. The nasal /m/ is slightly shorter in duration than the foregoing sounds, but it is quite clearly longer than the other voiced consonants /v n d r l/.

The intrinsic duration of the consonants was seen to some extent to be dependent, among other things, on their point of articulation.[1] According to Eli-Fisher-Jørgensen (1964, p. 189), of the lax plosives /b d g/ read in nonsense words by Danish readers, the labial plosive is significantly the longest in dura-

[1] Fintoft (1961, p. 32) has not noticed that the point of articulation has any effect on the duration of consonants in his Norwegian material.

TABLE 6. Absolute and intrinsic ($\overline{X}_C/\overline{X}_{allC}$) consonant durations in different nonsense-word structures. The individual values are means of the durations in words spoken by ten informants.

	single consonant								double consonant	
	I		II		III		I-III			
medial	/pa—a/		/pa—an/		/pa—aa/		mean		/pa—a/	
consonant	msec	intr.	msec	intr.	msec	intr.	msec	intr.	msec	intr.
/p/	102	1.40	99	1.34	118	1.40	106	1.38	191	1.17
/t/	94	1.29	94	1.27	108	1.29	99	1.29	197	1.21
/k/	98	1.34	102	1.38	112	1.33	104	1.35	205	1.26
/s/	94	1.30	82	1.11	101	1.20	93	1.21	190	1.16
/h/	79	1.08	76	1.03	85	1.00	80	1.04	–	–
/m/	67	0.92	73	0.99	79	0.93	73	0.95	142	0.87
/n/	56	0.77	55	0.75	65	0.77	59	0.77	134	0.83
/ŋ/	–	–	–	–	–	–	–	–	146	0.90
/l/	48	0.66	52	0.70	52	0.62	51	0.66	124	0.76
/r/	45	0.61	50	0.67	61	0.73	52	0.68	125	0.77
/j/	83	1.14	88	1.19	100	1.18	90	1.17	–	–
/v/	57	0.79	60	0.82	69	0.82	62	0.81	–	–
/d/	51	0.70	53	0.72	62	0.74	55	0.72	168	1.03
means of all allophones	72.5		73.4		84.0		77.0		162.5	

tion, the dental and velar plosives are of about the same duration, even though the dental tends, in general, to be longer.

According to Meyer and Gombocz (1909, p. 154), labial plosives are longest and dental plosives shortest in both Hungarian and English. Ariste (1941, p. 10–13) says, in his description of the dialect of the Estonian archipelago, that the labials are the longest of both the plosives and the nasals, next come the dentals and the shortest of the plosives is the velar plosive /k/. The labial plosive is also the longest in the South Ostro-Bothnian dialect of Finnish, where, according to Laurosela (1922, p. 89–96), the respective durational order of single plosives and various single and of geminate nasals is as follows:

/p/	(129 msec)	m	(98 msec)	/mm/	(253 msec)
/k/	(111 msec)			/ŋŋ/	(245 msec)
/t/	(101 msec)	n	(66 msec)	/nn/	(215 msec)

According to the means of the geminate plosives calculated by Laurosela, the dental plosive is just as long as the labial plosive but the present author obtained a result of 232 msec for the average duration of /pp/ (using the sentence

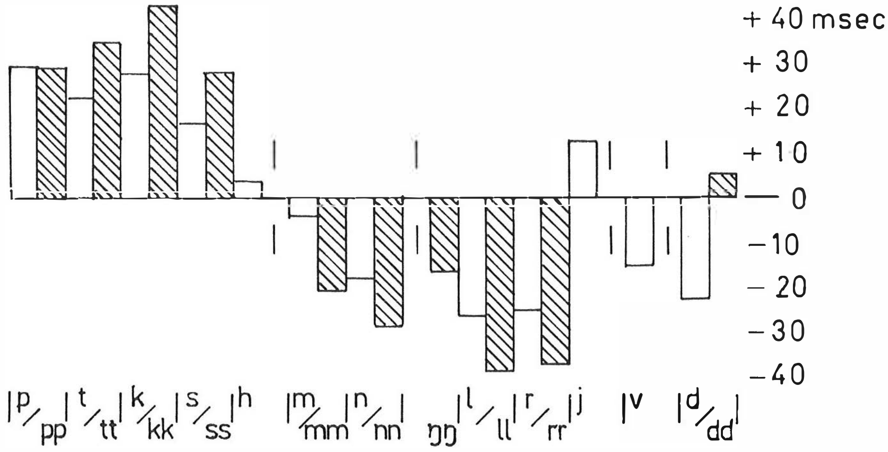


Figure 4. The durational differences of various intervocal single consonants (empty columns) and of double consonants (lined columns). The zero-level is the average mean of the compared consonants in each series (77 msec for single consonants, 163 msec for double consonants). The velar nasal /ŋ/ does not appear as an intervocal single consonant: the fricative /h/ and the semi-vowels /j/ and /v/ do not appear as double consonants. The figure is drawn according to table 6. The statistical significance of the differences is tested in table 7.

list published by Laurosela) and of 210 msec for the duration of the velar and dental geminate.[1] But the study made in Swedish by Elert (1965, p. 148) concerning the durations of consonants with various points of articulation is completely at variance with the results presented above. According to Elert, the velar consonants are longer than the dental and retroflex consonants and, moreover, the dental consonants are significantly longer than the labial consonants.

In the nonsense material of the present research read by native Finnish speakers, the dental consonants are shorter than the corresponding labial and velar consonants in all the groups of single consonants, even though the diffe-

[1] The scattered information of consonant durations in the Tornio dialect of Finnish given by Airila (1912, p. 56–58), on the other hand, denotes the priority of velar plosives in duration over the alveolars.

rence is only statistically significant between the durations of the single nasals m/n. The relationships between geminate consonants, on the other hand, seem complex: the order of double plosives is, starting with the longest /k t p/, but of the double nasals /ŋ m n/. None of these differences is, however, statistically significant. It must be noted that in the present comparisons the vowel on both sides of the consonants have always been an open back /a/. The durational interrelationships between consonants in different positions probably also depend on the nature of surrounding vowels.

TABLE 7. Comparison of the differences of durations between single consonants in nonsense-word structures CVCV (I), CVCVC (II) CVCVV (III), and double consonants in nonsense-word structure CVCCV. The level of significance of the difference is given as a percentage.

Consonants paired	single consonant I $\bar{d}$ (msec)	p	single consonant II $\bar{d}$ (msec)	p	single consonant III $\bar{d}$ (msec)	p	Mean of I–III (msec)	double consonants $\bar{d}$ (msec)	p
/p-t/	+8	non	+5	non	+10	non	+8	–7	non
/p-k/	+4	non	–3	non	+6	non	+2	–14	non
/p-s/	+8	20%	+17	2%	+17	10%	+14	+1	non
/p-h/	+23	0.1%	+23	0.1%	+33	1%	+26	–	
/p-m/	+35	0.1%	+26	0.1%	+39	0.1%	+33	+49	1%
/p-n/	+46	0.1%	+44	0.1%	+53	0.1%	+48	+57	0.5%
/p-ŋ/	–		–		–		–	+44	5%
/p-l/	+54	0.1%	+47	0.1%	+66	0.1%	+56	+67	0.5%
/p-r/	+57	0.1%	+49	0.1%	+57	0.1%	+54	+66	0.5%
/p-j/	+19	1%	+11	10%	+18	10%	+16	–	
/p-v/	+44	0.1%	+39	0.1%	+49	0.1%	+44	–	
/p-d/	+51	0.1%	+46	0.1%	+56	0.1%	+51	+23	non
/t-k/	–4	non	–8	non	–3	non	–5	–8	non
/t-s/	–1	non	+12	20%	+8	non	+6	+8	non
/t-h/	+15	2%	+18	1%	+24	2%	+19	–	
/t-m/	+27	0.1%	+21	20%	+30	0.1%	+26	+56	0.5%
/t-n/	+38	0.1%	+39	0.1%	+44	0.1%	+40	+63	0.5%
/t-ŋ/	–		–		–		–	+51	2%

Consonants paired	Single consonant I $\bar{d}$(msec)	p	II $\bar{d}$(msec)	p	III $\bar{d}$(msec)	p	Mean of I–III (msec)	double consonants $\bar{d}$(msec)	p
/t-l/	+45	0.1%	+42	0.1%	+56	0.1%	+48	+73	0.5%
/t-r/	+49	0.1%	+44	0.1%	+42	0.1%	+45	+72	0.1%
/t-j/	+11	20%	+6	non	+9	non	+9	–	
/t-v/	+36	0.1%	+34	0.1%	+46	0.1%	+39	–	
/t-d/	+43	0.1%	+41	0.1%	+46	0.1%	+43	+29	20%
/k-s/	+4	non	+20	1%	+11	10%	+12	+15	non
/k-h/	+19	0.1%	+26	0.1%	+27	0.5%	+24	–	
/k-m/	+31	0.1%	+29	0.1%	+33	0.1%	+31	+63	0.1%
/k-n/	+42	0.1%	+47	0.1%	+47	0.1%	+45	+71	0.1%
/k-ŋ/	–		–		–		–	+58	0.5%
/k-l/	+50	0.1%	+50	0.1%	+60	0.1%	+53	+81	0.1%
/k-r/	+53	0.1%	+52	0.1%	+51	0.1%	+52	+80	0.1%
/k-j/	+15	2%	+14	5%	+12	5%	+14	–	
/k-v/	+41	0.1%	+42	0.1%	+43	0.1%	+42	–	
/k-d/	+47	0.1%	+48	0.1%	+50	0.1%	+48	+37	5%
/s-h/	+16	0.1%	+6	non	+16	5%	+13	–	
/s-m/	+27	0.1%	+9	20%	+22	0.1%	+19	+48	0.5%
/s-n/	+39	0.1%	+27	0.1%	+27	0.1%	+31	+56	0.5%
/s-ŋ/	–		–		–		–	+43	2%
/s-l	+46	0.1%	+30	0.1%	+49	0.1%	+42	+66	0.1%
/s-r/	+50	0.1%	+32	0.1%	+40	0.1%	+41	+65	0.1%
/s-j/	+11	10%	–6	non	+1	non	+2	–	
/s-v/	+35	0.1%	+22	0.5%	+32	0.1%	+30	–	
/s-d/	+43	0.1%	+29	1%	+39	0.1%	+34	+22	20%
/h-m/	+12	0.5%	+3	non	+6	non	+7	–	
/h-n/	+23	10%	+21	0.1%	+20	1%	+21	–	
/h-l/	+31	0.1%	+24	0.1%	+33	0.1%	+29	–	
/h-r/	+34	0.1%	+26	0.1%	+24	5%	+28	–	
/h-j/	–5	non	–112	2%	–5	5%	– 11	–	
/h-v/	+22	0.1%	+16	1%	+16	10%	+18	–	
/h-d/	+28	0.1%	+23	2%	+23	3%	+25	–	
/m-n/	+12	10%	+18	0.1%	+14	0.5%	+15	+8	non

Consooants paired	single consonant I d̄(msec)	p	II d̄(msec)	p	III d̄ (msec)	p	Mean of I–III (msec)	double consonants d̄(msec)	p
/m-ŋ/	–		–		–		–	–5	non
/m-l/	+19	0.1%	+21	0.1%	+27	0.1%	+22	+18	non
/m-r/	+23	0.5%	+23	0.5%	+18	5%	+21	+17	non
/m-j/	–16	2%	–15	0.5%	–21	0.1%	–17	–	
/m-v/	+10	10%	+13	5%	+10	10%	+11	–	
/m-d/	+16	1%	+20	5%	+17	2%	+18	–27	20%
/n-ŋ/	–		–		–		–	–12	non
/n-l/	+8	20%	+4	non	+13	5%	+8	+10	non
/n-r/	+11	20%	+6	non	+4	non	+7	+9	non
/n-j/	–28	0.1%	–33	0.1%	–35	0.1%	–32	–	
/n-v/	–2	non	–5	non	–5	non	–4	–	
/n-d/	+5	non	+2	non	+3	non	+	–34	10%
/ŋ-l/	–		–		–		–	+22	non
/ŋ-r/								+21	non
/ŋ-d/								–22	non
/l-r/	+4	non	+1	non	–9	non	–1		non
/l-j/	–35	0.1%	–36	0.1%	48	0.1%	–40	–	
/l-v/	–9	10%	–9	20%	–17	2%	–12	–	
/l-d/	–3	non	–2	non	–10	20%	–5	–44	5%
/r-j/	–39	0.1%	–38	0.1%	–39	0.1%	–39	–	
/r-v/	–13	10%	–11	20%	–8	non	–11	–	
/r-d/	–7	non	–4	non	–1	non	–4	–44	5%
/j-v/	+26	0.1%	+28	0.1%	+21	0.1%	+25	–	
/j-d/	+32	01.%	+35	0.1%	+38	0.1%	+35	–	
/v-d/	+6	non	+7	non	+7	non	+7	–	

3.6.3. The effects of adjacent sound on sound duration

3.6.3.1. The effect of a preceding and following consonant on the duration of a vowel.

3.6.3.1.1. The effect of word-initial consonants on vowel duration

There is conflicting information from various languages on the effect of a word-initial consonant on the duration of the following vowel. For example, Peterson & Lehiste (1960, p. 701) show that in English '. . .the influence of an initial consonant on the duration of the syllable nucleus followed no simple regular pattern.' Fintoft came to the same conclusion in regard to Norwegian (1961, p. 26) as did Meyer and Gombocz (1909, p. 135) in Hungarian. Elert has not treated the question in the light of Swedish material in his study on Swedish quantity, but he refers to several studies in which there has been no proof that the initial consonant has an effect on the duration of the following vowel (Elert 1965, p. 132–140). On the other hand, research in Danish has revealed that the vowels following the Danish consonants /p t k f s h/ are slightly shorter than after other consonants in the language. (Fischer-Jørgensen 1964, p. 185–187). According to Laurosela (1922, p. 230), the vowels following /s/ and /h/ in the South Ostro-Bothnian dialect of Finnish are often considerably shorter than vowels following other consonants.

In the present study there is no series of test-words where initial consonants appear in identical surroundings, so the effect of initial consonants cannot be systematically examined on the basis of this material. In the test words of the sentence list, there are, however, a few series of word structures in which the difference between the words is only in their initial consonant (*takan*, *vakan*, *lakan*/ *takkaa*, *vakkaa*, *lakkaa* etc.; *tapa*, *rapa*, *lapa*, *napa*/ *tapaa*, *rapaa*, *lapaa*, *napaa* etc,; *salaa*, *halaa*/ *salata*, *halata* etc.) In these words, however, the sentence frame is not the same and in some cases consonant clusters appear on the word boundary: these things could affect the duration of the vowels. The differences between vowel durations, however, in the separate minimal series, show that there might be a certain regularity in the influence of different word-initial consonants on the following vowels: for example, the vowel /a/ is, in all cases, shorter after /h/ than after /s/ (the shortening is, on average, 20 %, 18 msec), just as a vowel seems to be slightly shorter after plosives than after nasals or liquids. But on the basis of a few scattered observations it is not possible to establish whether the effect of word-initial consonants differs in, for example, magnitude from the effect of an intervocal consonant on the subsequent vowel. It is obvious, however, that the initial consonants affect the duration of the following vowel in a similar manner to word-medial consonants. Eli Fischer-Jørgensen also reached the same conclusion (1964, p. 197).[1]

3.6.3.1.2. The effect on vowel duration of a consonant following a vowel

As has been shown with reference to various languages, the duration of a vowel is not only affected by its own phonetic characteristics, but also by the quality of the following consonant. Particular interest has been shown in this feature in English, where a vowel preceding a voiceless consonant is noticeably shorter in duration than preceding a voiced consonant. The durational relations between English vowels and, in particular, the effect of consonant voice have been examined by Denes 1955, Belasco 1953, 1958, Zimmerman & Sapon 1958, Peterson & Lehiste 1960, House 1961, Sharf 1962, Delattre 1962, Wiik 1965, to name but a few. For example, according to Sharf (1962, p. 29) the duration ratio of vowels preceding voiceless and voiced stops is about 3:4 for vowels preceding /p b/ and about 4:5 for vowels preceding /kg/. The ratio of vowel before voiceless consonant to vowel before voiced consonant reported by Peterson & Lehiste (1960, p. 702) is approximately 2:3. According to Wiik (1965, p. 116) the durational difference between vowels can be even greater: in Wiik's material, vowels preceding voiced consonants are 82 % longer than those preceding voiceless consonants. The effect of consonant voice on the preceding vowel is not only a characteristic phonetic feature in English, but also a cue which alone distinguishes a voiceless from a voiced consonant in perception (Denes 1955). Two partly conflicting methods have been used to explain this phenomenon. On the one hand, it has been said that the shortening effect of a voiceless consonant on the preceding vowel compared to the corresponding voiced consonant is a universal phonetic phenomenon. In English the shortening of a vowel preceding a voiceless consonant would also be the automatic, physiologically con-

[1] The shortening effest of *h* has also been observed by Laurosela (1922, p. 230) and Eli Fischer-Jørgensen (1964). The shortening effect of plosives has also been illustrated in Estonian. The shortening of vowels is strongest both absolutely and percentually in the I quantity degree words and smallest in the III quantity degree words (Liiv 1961, p. 422). In Hungarian there is, according to Magdics (1961, p. 43), a difference in the influence of various plosives on the duration of the preceding vowel so that, for instance, the ratio of a vowel preceding /t/ to a vowel preceding /k/ would be as much as 1:0.8.

ditioned anticipation of greater effort needed for the subsequent /tense/ consonant (Delattre 1962; Belasco 1958). On the other hand, it has been suggested that variations in vowel duration may be conditioned less by the articulatory features of consonants than by the linguistic structure of English. According to this idea, consonant voice or articulation power may have some universal phonetic influence of its own on the duration of the preceding vowel, but in English the variation in vowel duration preceding a voiced or voiceless consonant is learned and intentional and forms an essential cue of the tense/lax distinction between consonants (Zimmerman & Sapon 1958; House 1961). According to the latter idea, it is to be assumed that different vowel timing should not be used in all other languages as a phonetic cue of consonant voice and that the effect of consonant voice on the duration of a vowel would be markedly smaller. Obviously the effect of voice or of lack of voice in a consonant on the duration of the preceding vowel is noticeably smaller in many languages.

In Swedish the lengthening of a vowel preceding a voiced plosive (in the pairs p/b, t/d) is only 9–17 msec. which is statistically only almost significant in the long vowel series. In the short vowel series the durational differences are greater than in the long vowel series and are also statistically strongly significant (Elert 1965, p. 134). The lengthening effect of a voiced consonant following a vowel on the duration of the vowel is very slight in Hungarian: in the short vowels it is 15 %, in the long vowels only 6 % (Meyer & Gombocz 1909, p. 136). A first syllable vowel in Lappish is on average 13.8 % longer before a voiced than before a wholly or partly voiceless consonant (Äimä 1914, p. 136).

The effect of consonant voice on the duration of a preceding vowel in Finnish cannot be measured by minimal word pairs as, for example, in English or Swedish, because consonant voice is not a phonemically distinctive feature in Finnish. Finnish does not have the phoneme distinctions p/b, k/g, s/z etc. and the only so-called lax consonant having a voiceless counterpart is /d/. Instead of voice the distinctive feature of /d/ may be seen as a more retracted point of articulation than /t/ (Sovijärvi 1957, p. 317; Wiik 1965, p. 24). The other possible voiced/voiceless pairs are quite marginal within the phonemic structure of Finnish because the voiced plosives /b g/ only appear in new unabsorbed loan words and, apart from loan words, the fricative /f/ only appears in some dialects in the west of Finland.

Wiik (1965, p. 117), in order to have an object for comparison with the corresponding phenomenon in English, measured vowel duration before the voiceless consonants /ptks/ and before the voiced consonants /d v/. Vowels were 10-16 msec longer before the voiced than before the voiceless consonants: the lengthening of vowels before voiced consonants was 17 % in the single vowels and 4 % in the double ones.

In the list of nonsense words used in this study both intervocal /t/ and intervocal /d/ are presented in four different bisyllabic word structures. In all of these the first syllable single vowel is longer before /d/ than before /t/, the lengthening being on average 11 msec (15 %). But intervocal /d/ also has exactly the same effect on the duration of the following vowel: in the same bisyllabic group the vowel following /d/ is also 15 % longer than the vowel following /t/, the absolute lengthening being on average 17 msec.[1] In the meaningful sentence group, there are only a few word pairs which could be used in an evaluation of the effect of /d/: in the pairs *mätän/ mädät*, *mutan/ mudan* (N together 20) the duration of the first vowel is 9 % longer in front of /d/ (6 msec) and only 3 % longer after /d/ than the corresponding vowels adjacent to /t/. In the pair *mätän/ mädät*, the difference in final consonant also affects the duration of V_2; yet even in the pair *mutan/ mudan* the duration of V_2 after /d/ is only 5 % longer. In the only minimal pair where V_1 is double, *muutan/ muudan* (N= 10) the lengthening of the first vowel in the word containing *d* is 17 % and the lengthening of the second vowel in the same word is 22 %.

Although a vowel is generally slightly longer next to the phonetically voiced plosive /d/ in Finnish than next to a voiceless /t/, there are no grounds for concluding that this lengthening is caused expressly by the presence of voice. The pair t/d is by no means a special case in investigations into the reasons for vowel differences in relation to various consonants. The following tendency observed in connection with other consonants, may also be found here: the longer the intrinsic duration of a consonant, the more it shortens the preceding vowel, and vice versa.

The effects of intervocal consonants on preceding (blank bars) and following (black bars) /a/ -vowels is presented graphically in figure 5. Nonsense word series of the columns I, II and III of table 6 were used as comparison structures in which each different consonant appeared in the same position in an identical word frame. The number of cases measured for each position is therefore 30. Vowels preceding consonants are all single, vowels following consonants are

[1] In Russian, too, a clear lengthening effect of consonant voice on both preceding and following vowel is indicated by Bolla (1968).

phonetically half-long single in series I and II and in series III they are double. The average duration (80 msec) of all V_1 segments is shown as zero in the figure as is the average duration (138 msec) of all V_2 segments. The bars show the average deviation from the total mean value in milliseconds of the allophones of /a/ appearing in the V_1 and V_2 position alongside each consonant.

According to the figure, all the voiceless plosives /p t k/ as well as the semivowel /j/ appear to have a shortening effect on the preceding and following vowel. The fricative /h/ seems to have a very strong shortening effect on the

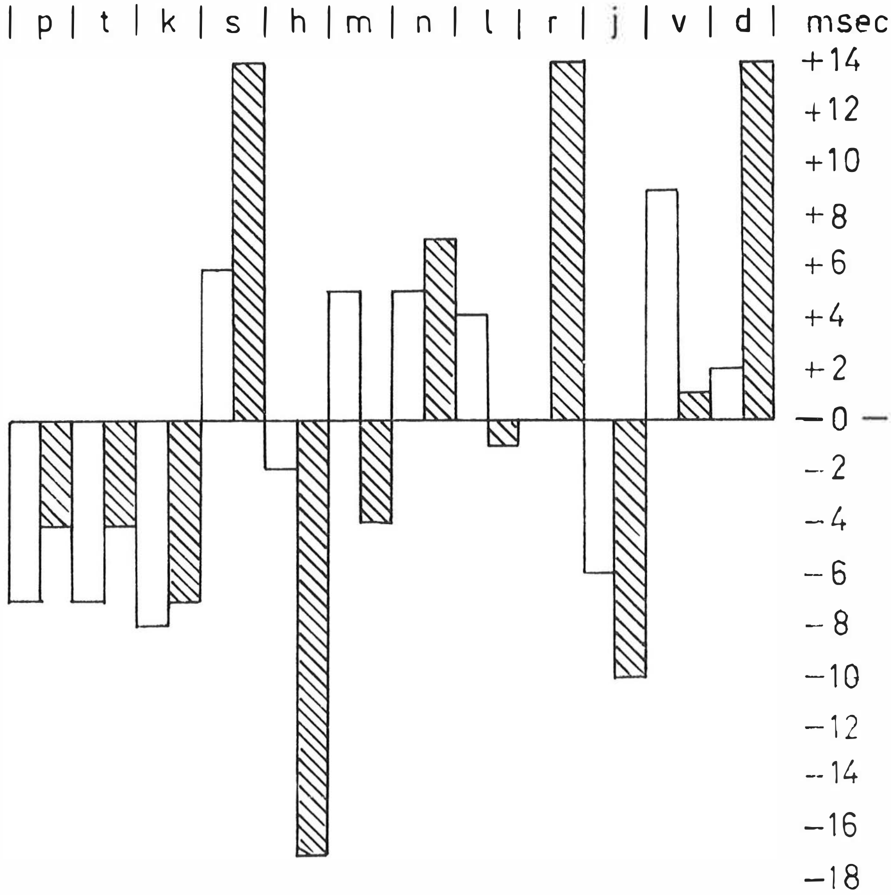


Figure 5. The effect of the different inter-vocal consonants on the duration of the preceding and following /a/ vowel. The mean duration of different /a/ allophones was taken as the zero-level in both positions. The white columns indicate the effect of the consonant on the preceding vowel, the lined columns the effect of the consonant on the following vowel.

following vowel, but to have less of an effect on the preceding vowel. On the other hand the effect of /s r d/ is to give more duration to the vowel following the consonant. The deviations from the total mean value caused by different consonants are absolutely and relatively fairly small: the absolute deviation is in all cases less than 20 msec and the greatest deviation (shortening of V_2 by /h/) is only 12 % of the average vowel duration. On the basis of the results, it cannot be said that the effect of a consonant on the preceding vowel is greater than its effect on the following vowel or vice versa: roughly speaking, the effect of a consonant on the preceding and its effect on the following consonant are of the same kind. Apart from the fricative /s/ which lengthened both the preceding and the following vowels, all consonants which shortened vowels, /p t k h j/, are the consonants of longest intrinsic duration.

A change in the segmentation procedure of the semi-vowel /j/ would, however, have affected its position in the figure: if the VC- and CV-boundaries of *j* had been determined nearer to /j/'s short steady state, instead of at the steepest point of the formant transitions, its own duration would have been shortened and at the same time its shortening effect on the surrounding vowels would have been reduced. Correspondingly, if the CV-boundary of plosives was fixed at the centre of the explosive burst, their shortening effect on the following vowel would be reduced. The voiced plosive /d/ is about equal in influence to the consonants /m n l r v/. Similarly, the effect of Danish /b d g/ on the preceding and following vowels has been shown to be the same as the effect of /m n l v/ (Fischer-Jørgensen 1964, p. 187; for English see also Peterson & Lehiste 1970, p. 700, 702, for Russian Bolla 1968).

3.6.3.1.3. The effect of the final consonant of a word on the durations of the preceding sound segments

As an illustration of the effect of open and closed final syllables, four different cases are compared in table 8: one of the comparison words ends in a vowel, the other has a consonant added to the identical vowel ending (e.g. *tapa* 'manner' / *tapan* 'I kill': *mättää* '(he) scoops'/ *mättään* 'the hillock's': *tuuppaa* '(he) shoves' / *tuuppaat* '(you) shove': *nasta* 'pin' / *nastan* 'the pin's').[1]

[1] It is particularly interesting to explain the effect of open/closed final syllable in a comparison between the phoneme development of Finnish and Estonian (cf. words of second and third degree of quantity in Estonian and their Finnish equivalents; est. vil:· /fin. *villa*; est. vil:a /fin. *villan*; est vil:·a /fin *villaa*<*villa-a*; est kan:· / fin *kannu*; est kan:u /fin *kannun*; est kan:·u / fin *kannu-a*: est my:·r / fin *muuri*; est my:ri / fin *muurin*; est my:·ri / fin *muuri-a*).

TABLE 8. Comparison of segmental durations in four word-structure pairs ending in a vowel and in a consonant. Differences of means are presented in milliseconds and the statistical significance of differences (according to t-tests) by asterisks. The number of compared pairs is 240 in the first set, 170 in the second and 50 in the third and fourth set.

Compared structures
CVCV/CVCVC

Sound segment	Structure CVCV	Structure CVCVC	Difference of means
C_1	72	67	+5*
V_1	66	63	+3*
C_2	75	71	+4*
V_2	104	100	+4***

Compared structures
CVCCVV/CVCCVVC

Sound segment	Structure CVCCVV	Structure CVCCVVC	Difference of means
C_1	65	73	–8**
V_1	77	78	–1 non
C_2	194	189	+5 non
V_2	134	124	+10

Compared structures
CVVCCVV/CVVCCVVC

Sound segment	Structure CVVCCVV	Structure CVVCCVVC	Difference of means
C^1	77	69	+8 (p<10%)
V^1	148	145	+3 non
C^2	181	178	+3 non
V^2	139	125	+14****

Compared structures
CVC_xC_yV/CVC_xC_yVC

Sound segment	Structure CVC_xC_yV	Structure CVC_xC_yVC	Difference of means
C_1	65	78	-13^{****}
V_1	83	86	-3^{non}
C_2	143	160	-17^{***}
V_2	46	59	-13^{****}
C_{2x}	76	86	-10^{*}
C_{2y}	66	73	-7^{non} ($p<10\%$)

The results of the comparison do not give a consistent picture of the effect of a final consonant. In the first three comparisons the final vowel is significantly or strongly significantly longer in a word ending in a vowel than in a word ending in a consonant: an open or closed second syllable has had no significant effect in these word pairs on the segments preceding the second syllable. However, in the fourth comparison, where all the word pairs are of the $CVC_xC_yV(C)$ type, the second vowel in a word ending in a consonant is very significantly longer than the final vowel of an open syllable. The effect of an open or closed final syllable may also be studied with the help of the nonsense-word series (columns I and II in table 6). Minimal word pairs CVCV and CVCVC appear here in completely identical sentence frames, which is not the case with the group of meaningful words. In each series of nonsense words the number of cases measured is 120. The average durations of segments are as follows:

	CVCV	CVCVC
C_1	107 msec	107 msec
V_1	79 »	79 »
C_2	73 »	73 »
V_2	115 »	124 »
C_3	–	78 »

Although the significance of the difference in means between the sounds is not statistically tested, the difference between the structures is obvious: the initial segments are identical in their duration while the vowel of the second syllable is longer in duration in a closed than in an open syllable. The effect of the open syllable is therefore the reverse of the effect of the corresponding open syllable in the majority of meaningful words. The reason for this discrepancy may be as follows: in both the comparison between CVC_xC_yV and CVC_xC_yVC and the nonsense word comparison, the nasal /n/ is the final consonant in all words ending in a closed syllable and this has a lengthening effect on the preceding vowel. On the other hand, in the rest of the comparisons in table 8 both nasals and plosives appear as final consonants, the latter having strong shortening effect on the preceding vowel. If, for example, the means of second vowel duration in CVCVC words ending in /t/ and CVCVC words ending in /n/ in the CVCV/CVCVC comparison series are calculated separately it can be seen that the former is slightly shorter (94 msec), the latter sightly longer (107 msec) than the mean of the V_2 of all CVCV words (104 msec). In the series of meaningful sentences the results of the comparison might also be influences by a different kind of preceding and following phonetic environment: this would explain why duration of the initial consonant turned out to be different in words ending in vowels and consonants. As a summary of these contradictory comparison results, one could, however, conclude that an open or closed final syllable has no effect as such on the durational pattern of a word. The final consonant of a word has its own intrinsic effect on the neighbouring vowels: in front of certain final consonants the vowel is shorter than the final vowel of an open syllable, in front of some others it is about the same duration and there are probably some cases where it is longer than the corresponding open syllable vowel. Even though this intrinsic effect of the final consonant is fairly small, word structures ending in vowels and consonants are dealt with separately in the following data.[1]

[1] The effect of the durational pattern of word structure is also related to the problem of the final morpheme {x} which has occupied students of the Finnish language (x= the doubling of the initial consonant of a following morpheme): is the imperative *anna* 'give', for example, in some way phonetically different from the girl's name *Anna?* Because not even the phonetically realized final consonant of a word changes the durations of preceding segments significantly, there is no reason to suppose that this is the case with a phonetically and phonemically unmanifested 'deep structure' final consonant. Within an utterance, however, the phonemic (and phonetical) representative of {x} influences the duration of the preceding vowel in the same way as any other consonant. There is no reason to assume, for example, that the utterance *Anna nuolet* /annannuolet/ 'give the arrows' should be phonetically different from the utterance *Annan nuolet* /annannuolet/ 'Anna's arrows' spoken with the same intonation pattern.

3.6.3.2. The effect of vowels on the duration of consonants

Just as different consonants can exert a different influense on the duration of vowels preceding and following them, it may be possible to think of different vowels as having an effect on the duration of adjacent consonants. Elert (1965, p. 150–154) has shown that in Swedish consonants following intrinsically longer open vowels are shorter than following intrinsically shorter closed vowels. Lip-rounding in the production of vowels is apt to increase the durational difference between consonants. According to Meyer & Gombocz (1909, p. 150), the plosives *k* and *t* are 20 % longer in duration after closed vowels than after open in Hungarian. On the other hand, Eli Fischer-Jørgensen (1964) has not discovered any consistency in the effects of preceding vowel on the duration of the closure in her Danish material. In the present study there is no sufficiently coherent nonsense-word list for the examination of the effects of different vowels on the durations of different consonants. However, some information about the effects of different vowels on the duration of the labial plosive /p/ was obtained incidentally from the nonsense-word lists which were meant to throw light on the durations of various vowels in an identical consonant frame. The bilabial plosive is slightly longer next to a rounded vowel than next to an unrounded vowel. Next to illabial vowels the mean duration of $/p/_1$ in the structure CVVCV is 114 msec, and the duration of $/p/_2$ 85 msec: next to labial vowels the corresponding durations are 126 msec and 95 msec. The lengthening caused by a labial vowel in the C_1 segment is 12 msec (11 %) and in the C_2 segment 10 msec (12%). In the structure CVCCVV the duration of $/p/_1$ is 99 msec in an illabial environment and the duration of $/pp/_2$ is 232 msec, while in a labial environment the respective durations are 105 msec and 238 msec. The lengthening of the C_1 segment of the structure CVCCVV caused by labial vowels is 6 msec (6 %) and of the C_2 segment also 6 msec (3 %). The effect of the vowel environment on a long consonant seems, according to this, to be both absolutely and relatively less than its effect on a short consonant. The effect of a labial vowel on the duration of consonants other than /p/ cannot be clearly illustrated because in the nonsense-word series only the plosive /p/ was used. On the basis of a few comparison structures in the sentence list (e.g. *mätä/muta/mutu* etc.) it seems that round vowels tend to lengthen other consonants besides the labials. It also seems likely that a consonant will be longer if it has a labial vowel on both sides than in the sequences $V_{ill}CV_{lab.}$ or $V_{lab}CV_{ill}$.

However, the material is far too limited for a statistical illustration of this tendency. The influence of vowel openness does not, on the basis of the nonsence word lists, seem as consistent and clear as the influence of lip-rounding. The plosive /p/, however, is on average slightly shorter next to the open illabial vowels /a/ and /ä/ than next to the more close illabial vowels /e/ and /i/.

3.7. *Durations of sound segments in various phoneme syntagms*

3.7.1. Durational distinction between short and long vowel quantity

In an examination of the objective vowel durations representing short and long vowels in different languages, attention should be paid to the level and magnitude at which the different quantities or subjective pronunciation times of sound segments are used in the structure of the language, and to the extent to which the long vowel/short vowel distinction is based on the durational differences between sounds or on other phonetical differences (see chapter 1.2). In the Germanic languages, such as German and the Scandinavian languages (Swedish, Danish and Norwegian) and English, where the long vowel/short vowel distinction only appears in word-stressed syllables, quality differences are also related to the durational differences between vowels. Apart from the difference in vowel duration, the distinction between two vowel categories in these languages is also carried by their different timbre features. The second group of languages is formed by those in which the long vowel/short vowel distinction is essentially based on the different durations between sounds and in which no noticeable quality difference has been observed to contribute to the long/short distinction between vowels. The languages of this group could be called quantity languages in their 'autonomous' use of the duration parametre (cf Ravila 1961). However, it must be supposed that quality differences are also related to the durational differences between vowels in all languages: generally, vowels which are long in duration are farther from the neutral position than those which are short. For example, long close vowels are closer in Finnish than short close ones (Sovijärvi 1938, p. 24). Similarly, it can be seen on Wiik's (1965, p. 59) formant chart of Finnish vowels that the average formant positions of single vowels are clearly situated nearer to the centre of the figure (F_1 500 cps, F_2 1500 cps) as compared to the average formant positions of the corresponding double vowels. In certain cases, the absolute differences between formant positions are even greater than between the nearest phonetically equivalent tense vowel/lax vowel distinctions in English. The decisive difference between these languages, however, is that a Finnish listener does not make use of the quality difference between vowels as a cue to the phonemic length of the vowel segment as an English listener does.

Elert (1965, p. 110) has grouped various languages into three classes according to the kind of position vowel quantity has in the phonemic system of each language: languages in which only a stressed vowel carries information about quantity belong to the first class (e.g. Danish, Dutch, German[1]), languages in which both vowels and consonants appear independently in different quantity degrees (e.g. Finnish, Hungarian, Czech) are in the second class, and the last class contains languages in which there is a correlation, at least in stressed syllables, between vowels and following consonants so that long vowels are followed by short consonants and vice versa (e.g. Icelandic, Norwegian, Swedish). Accordingly it would be natural to suppose that 1) the phonetic duration distinction between short and long sounds is greater in languages in which the domain of durational distinction consists of only one sound segment than in those where the contrast between a vowel and the following consonant also gives some information about the vowel length and 2) the durational difference between vowels is greater in languages in which the quantity difference is based, above all, on the durational difference between sounds than in languages in which the distinction is based both on different duration and on different quality.

These phenomena illustrated in various languages vary considerably depending on the research method used, but one can state that the above suppositions more or less hold true. According to Elert (1965, p. 113), the V/V: ratio in Swedish varies in different vowel distinctions from 62 % to 77 % and is on

[1] The distinction between Elert's first and second language class seems, at least in some cases, to be partly artifical: in the form of Swedish spoken in Finland, there is not, according to the observations of the author's Finnish-Swedish colleagues, any subjectively perceptible difference between duration of obstruents following long and following short vowels, e.g. visa/vissa, heta/hetta in Finnish-Swedish [vi:s:a/vis:a/he:t:a/het:a] (However, a recent test made in the phonetic laboratory of Jyväskylä University showed the mean ratio 1:1.36 between the obstruents following long and short vowels). On the other hand Delattre and Hohenberg (1968, p. 388) have shown that in German, in which according to Elert's classification a stressed vowel alone contributites to the phonemic quantity of the vowel, the influence of the duration of a subsequent consonant is not at all a negligible cue for the tense/lax (=long/short) distinction of a vowel, although the duration of a vowel and the colour of a vowel are, more often, stronger cues of phonemic length.

average 69 %, i.e. the duration of a short vowel is on average 69 % of the duration of the corresponding long vowel. The corresponding V/V: ratio in Norwegian, according to Fintoft (1961, p. 24) is 53 %, but according to Vanvik (1966, p. 56-58) the ratio in one syllable words in the Trondheim dialect of Norwegian is 1:2,4 – 1:2,6. The V/V: ratio in Danish is 50.5 % (Fischer-Jørgensen 1955, p. 43). Information concerning the ratio between long and short vowels in German varies considerably (see Delattre & Hohenberg 1968, p. 369) The average V/V: ratios in continuous speech, provided by the so-called phonometric school, vary from 51 % to 65 % (Zwirner & Zwirner 1937, Maack 1951), which is relatively close to the tense/lax durational ratio between vowels in an unstressed position of Delattre and Hohenberg, 1.52:1. In Dutch a long and short vowel are on average in a 1:1.6 ratio to each other (Kaiser 1964, p. 246). In English phonemically long vowels are on average 77 % longer than short ones, according to Wiik (1965, p. 114).

In Czech, which belong to Elert's second class, a short and a long vowel are in a 1:2 ratio to each other (Kaiser 1964, p. 246). According to Abramson (1962, p. 93), in Thai, in which the vowels, but not the consonants, can be either single or double, the double vowels are 2 to 3.5 times as long as the single vowels in an analogous environment. In running discourse the average ratio of double vowels to single vowels is 2.57.

Of the Finno-Ugric languages, Hungarian has an average V/V: ratio of 1:2 (Laziczius 1966, p. 75). In Lappish, in which Äimä distinguishes four quantity degrees, the average durations of vowels in the various quantity degrees are, in milliseconds, 137, 171, 304 and 382 (Äimä 1914, p. 144). In Estonian the vowels of the first, second and third quantity degrees have the following ratio to each other: 1 : 2, 19 : 3.53 and the vowels of the second and third degrees 1 : 1.61 (Liiv 1961, p. 480).

According to Donner's (1912, p. 10–14) figures, the ratio between short and long first syllable vowels in the Salmi dialect of the Karelian language can be calculated as 1 : 2.3 and the corresponding second syllable ratio as 1 : 2.1.

In the South Ostro-Bothnian dialect of Finnish, first syllable short and long vowels are in a 1 : 2.3 ratio to each other, and when all syllable positions are accounted for, the ratio is 1 : 2.45 (Laurosela 1922, p. 212–213). Palomaa (1922, p. 213) shows that in Standard Finnish the ratios between single and double vowel durations are: in the first syllable 1 : 2.08, in the second syllable 1 : 2.3, in the third syllable 1 : 2.5. According to Wiik's wider material read in a sentence frame, the durational ratio between a single vowel and a double vowel in the first syllable of Finnish bisyllabic words is on average 1 : 2.3 (Wiik 1965, p. 115). Penttilä (1963, p. 21) reports the ratio of short vowels and consonants to the corresponding long vowels and consonants in Finnish as 1 : 2.5 – 1 : 3.

In the present study, the average duration (N=2698) of all first syllable single vowels in words read in the meaningful sentence frames is 70 msec and the average duration of all first syllable double vowels (N=874) is 153 msec. First syllable single and double vowels are, therefore, in a 1 : 2.2 ratio to each other. The average ratio corresponds very well to the ratio obtained by Wiik. There is no reason to calculate the average ratio of the single and double vowels of subsequent syllables because it varies quite considerably from one word structure to another depending on the appearance of the half-long vowel in place of the functionally single vowel (see chapter 3.8). It can be said that in continuous speech a first syllable double vowel is at least twice as long as a single vowel in the corresponding position. This ratio is also supported by the minimal distinction between the durations of single and double vowels shown by perception tests, i.e. the durational ratio between the longest vowel perceived as a single vowel and the shortest vowel perceived as a double vowel (about 1 :1.9; see chapter 4). In slow or careful speech of moderate tempo, as well as in strong sentence stress positions, double, phonetically long vowels lengthen slightly more than single vowels so that the ratio between sound segment durations representing single and double vowels changes, but the variation is nevertheless generally below 1:3. The V/VV distinction also varies in the first syllable, however, depending on what kind of structure the first syllable single or double vowel appears in. In table 9. the average durations of the single and double vowels of various stuctures in the sentence list and their ratios have been brought together.

It is evident that the duration of a first syllable single vowel varies very little in different structures. It should be noted that it is longest where it is followed by a sequence of double consonant and double vowel, and shortest in structures where it is followed by a sequence of single consonant and single vowel. The first syllable double vowel is longest in structures where it is followed by a single consonant and a double vowel and about the same in duration if it is followed by a single consonant and single vowel or by a double consonant. The duration of a double vowel is in all cases at least twice as long as a single vowel

TABLE 9. The durational difference in the V/VV distinction in various word-positions

compared word pair	Mean durations /V/ msec	/VV/ msec	V/VV ratio	
CVCV/CVVCV	63	143	1 : 2.27	1. syllable
CVCVC/CVVCVC	63	148	1 : 2.35	
CVCVV/CVVCVV	67	168	1 : 2.50	
CVCVVC/CVVCVVC	68	176	1 : 2.59	
CVCCV/CVVCCV	66	132	1 : 2.00	
CVCCVV/CVVCCVV	65	143	1 : 2.20	
CVCCVVC/CVVCCVVC	72	145	1 : 2.02	
CVCV/CVCVV	105	145	1 : 1.38	2. syllable
CVCVC/CVCVVC	102	151	1 : 1.48	
CVCCV/CVCCVV	48	134	1 : 2.80	
CVVCV/CVVCVV	48	135	1 : 2.81	
CVVCCV/CVVCCVV	43	137	1 : 3.19	
CVCVVCV/CVCVVCVV	53	147	1 : 2.77	3. syllable
CVCVCCV/CVCVCCVV	40	115	1 : 2.86	
CVVCVVCV/CVVCVVCVV	61	132	1 : 2.16	

in the corresponding position. The difference is slightly smaller before a double consonant than before a single consonant. If the second group in table 9 is disregarded, it can be shown that a second and third syllable single vowel is on average shorter than a first syllable single vowel and that, correspondingly, the ratio between a single and double vowel is also greater. In one structure pair the double vowel is as much as three times as long as a single vowel in the corresponding position.

If all the durations of single vowels in table 9, measured in different structures and different positions, are examined, the structures of the second group CVCV(C)/CVCVV(C) differ very clearly from the others: the second syllable single vowel following a short initial syllable is considerably longer in duration than when following a long initial syllable and, correspondingly, the ratio between single and double vowels is clearly smaller than in any other case – in both comparison pairs it is smaller than 1 : 1.5. In the nonsense word series in which different single consonants appeared between /a/ -vowels in the structures papa/papaa, pata/pataa etc. the second syllable V/VV ratio is about the same, 1 : 1.51. Wiik and Lehiste (1968, p. 571) have compared the durational ratios of the vowels in (C)VCVC, (C)VVCVC, (C)VCVVC and (C)VVCVVC words as read by informants from different parts of Finland, and have established that the informants may be divided into two groups according to whether they use the so-called half-long vowel in the second syllable of the (C)VCV(C) structure or a vowel duration which was only slightly greater than the duration of the vowel of the initial syllable. In the latter group the second syllable V/VV ratio, calculated from Wiik and Lehiste's table, between the structures (C)VCVC and (C)VCVVC is 1 :1.9 and in the former group it is 1:1.5. Therefore the V/VV ratio of this study corresponds very well to Wiik and Lehiste's ratio between 'half-long' single vowel and double vowel.

3.7.2. Durational C/CC distinction of consonants

3.7.2.1. Quantity and the Anschluss concept

The concept of contact (German 'Anschluss') frequently appears in various connections in descriptions of Finnish quantities. Hakulinen (1961, p. 21–22) establishes that contact between vowels and consonants is close (fester Anschluss) before a geminate consonant, but open (loser Anschluss) before a single consonant. Sadeniemi (1949, p, 34–39) even considers that contact, together with sound sonority, is the decisive factor in the definition of Finnish syllable boundaries. Wiik (1965, p. 118) has also made use of the concept of contact for the classification of factors influencing the duration of Finnish vowels: according to Wiik, the duration of a long vowel is longer if the contact between vowel and consonant is open (for example the [u:] vowel in the word *muuta* 'something else'), and shorter if the contact between vowel and consonant is close (for example the [u:] vowel in the word *muutta* 'without anything else'). Wiik also states that contact functions as a kind of boundary signal. 'In Finnish, a »loser Anschluss» between a vocoid and a contoid denotes a syllable boundary, while in a »fester Anschluss» the vocoid and the contoid belong to the same syllable'. The concept of contact is also used in the description of

Estonian quantities: the vowel representing the third degree of quantity is joined to the following consonant by open contact, whereas contact between a second degree of quantity vowel and the following consonant is close. Liiv (1961, p. 488) supposes that difference in contact could be one important perception cue fo distinguishing between the second and third degrees of quantity. However, the concept of contact seems very vague and one may ask what are the cause and effect in cases of different kinds of contact.

The Anschluss concept traditionally means the subjective impression that a consonant seems to break the vowel during its intensity maximum in close contact, whereas in the case of open contact the intensity of a vowel has already weakened when a consonant begins. v. Essen (1962) has tried to show the connection between the abatement of amplitude of vowel oscillation and the German contact distinction (e.g. *hüte/Hütte*). This was justly criticised, in the same connection, by Eli Fischer-Jørgensen: the differences between the intensity progress of long and short vowels can also be illustrated in languages where there is no fester/loser Anschluss distinction. According to Fischer-Jørgensen, the subjective impression of fester Anschluss in German may be considered a secondary consequence of vowel shortness and laxity (Ungespannheit).

In many cases one can even see that the abatement of intensity is quicker when the perceived contact is open and slower where it is close (see Fintoft 1961). It is easy to illustrate, by means of perception tests, that the subjective impression of contact can, without interfering with the intensity progress of sounds, be changed only by changing the durational ratio of a vowel to the following consonant: for example, by shortening the voiceless phase of the plosive in the word *matto* 'carpet', close contact becomes open; correspondingly, by shortening the voiceless phases of the plosive in *muutta* or by lengthening the voiceless phase of the plosive in *muuta*, the original impression of Anschluss changes. The same effect is obtained if the vowel's instead of the consonant's original duration is changed (cf Fliflet 1962, 1963). The perception tests seem to indicate that the durational ratio of a vowel to the following consonant determines Anschluss perception and that the difference in abatement of vowel amplitude does not contribute to the phenomenon at all. According to the present author's own observations, the Anschluss phenomenon is only related to the stressed syllables of a word. In the word pairs *kasata* 'to heap' /*kasatta* 'without a heap' and *kaipuuta* 'of longing'/ *kaipuutta* 'without longing' pronounced normally by stressing the first syllable, the difference in contact between the unstressed second syllable vowel and the following consonant cannot be observed. The observation of close contact may be related to the breaking of the dynamic contour (of melody and intensity) supporting the perception of stress. The longer the time or the earlier the consonant breaks the acoustic stress pattern produced by the sub-glottal exhalation pulse, the closer the contact between vowel and consonant is perceived to be.[1]

3.7.2.2. The concept of geminate consonants

As was shown in connection with Finnish phonemics (chapter 2), Finnish long consonants, traditionally called geminate consonants, are the complete equivalents of clusters of two different intervocal consonants: for example, the words *kassaa* 'of cash', *kattaa* '(he) covers' *kastaa* '(he) dips', *kantaa* '(he) carries' are structurally completely alike. Considering a long consonant as a double consonant, i.e. as a consonant cluster, is not, therefore, an arbitrary problem of interpretation.[2] Traditionally only a long consonant containing a syllable boundary can be called a geminate consonant (Malmberg 1967).[3] When the syllable boundary divides a geminate consonant in two, it can also be produced in reading as two separate parts (*kas-saa* just as *kas-taa*). This possibility has lead to the supposition that the geminate consonant would be characteristically double peaked in intensity (see Dieth 1950, p. 415–416). However, it has not been possible to observe any indication of the second intensity peak or any other kind of phonetic boundary between the members of a geminate on the acoustic level of analysis (v. Essen 1966, p. 168; for Hungarian geminates in

[1] About the phonetic manifestation of stress see Lieberman 1967; 1970.

[2] According to Ravila (1961, p. 348) the geminate concept is a difficult problem in phonemic theory since one should be able to define a syllable boundary in order even to talk about geminates. This is very likely true if geminates and the corresponding single consonants are examined in isolation instead of the distribution of all consonants.

[3] In his study entitled 'Die Quantität als phonetisch-phonologischer Begriff' (1944, p. 78) Malmberg sought to distinguish between phonemically non-distinctive long consonants (for example, Swedish *matta, kanna*) and 'genuine or proper geminates' which are phonemically always combinations of consonants (for example, Swedish *bordduk*, *takkrona*); however, on the syllabic level, the former can also be geminates in Swedish (Malmberg 1967, p. 153). Ariste (1953, p. 94–95) also relates the geminate concept to, in particular, distinctive function when he says that so-called geminates seem to appear in languages where phonological quantity exists. Ariste seems to think that in a long consonant which is to be regarded as a geminate the double syllabic character of the members and the pause between them (!) must also be audible by the ear; therefore, according to Ariste, the long consonants of Estonian are not, given these conditions, geminates.

partic. see Hegedüs 1959). It has only been established that geminate consonants are phonetically single long consonants whose longer duration is the only signal of the fact they are bisyllabic and contain two members. However, in her analysis of duration in Estonian, Lehiste (1960, p. 52; 1966, p. 17, 70) has observed spikes on the spectrogram within the gap of voiceless plosives, which, according to Lehiste, indicate the syllable boundary. Those strokes which resemble spikes observed within clusters of plosives whose members are produced with a differing articulatory process, correspond, according to the author, to a subjective kinesthetic impression of rearticulation of the plosive. Corresponding peaks are not to be seen in the Finnish material of the present study. It is possible – and even probable – that the members of word-medial geminate consonants belong on some deeper production level to separate units, but that hypothetically separate production commands assimilate in the articulation process in such a way that rearticulation of the latter part of the geminate does not even appear in the action of the articulatory muscles. The phonetic output is, in any case, just one long consonant segment without any observable 'boundary signals'. It should be noted that not even in the steady states of phonetically over-long consonants in the slow-speech words of this research can one observe any signs of rearticulation – no intensity spikes during plosives, no changes in intensity or quality during long fricatives. Rearticulation of the second member of a geminate is, however, possible and is even performed when the word containing the geminate consonant is read syllable by syllable.

In Finnish a phonetically long consonant does not necessarily always have to be a geminate consonant, i.e. divided in two by a syllable boundary. Since the syllable boundary in Finnish always occurs in front of a consonant preceding a vowel, a geminate can also, therefore, only appear in this position. In the completely possible expression [met:sä] =/metsä/ 'woods' the [t:] is not a geminate, but merely a 'long consonant' whose duration, being greater than normal, does not effect the syllable boundary in any way, and is not phonemically relevant.

The concept of syllable boundary also plays a certain part in the evaluation of the duration of geminate consonants and in the subjective segmentation of their members. When an intervocal single consonant is syllabilized together with the following vowel, it is logical to suppose that the additional duration in the geminate consonant compared to a single consonant in the same environment (e.g. *takka*/*taka*) has been added to the geminate before the single consonant and that when lengthening or shortening the geminate, this initial portion of the long consonant is lengthened or shortened.

This concept is also reflected in the set of symbols used in the Finno-Ugrian transcription system: a long consonant which the ear hears as shorter than a 'normal' double consonant is called in Finnish 'lyhytalkuinen' ('having a short beginning') and is represented by a 'short duration' symbol placed on the preceding consonant, for example [t̆t]. Similarly, a consonant which is in subjective duration longer than a normal geminate is represented by a 'half-long duration' diacritical mark placed on the preceding letter e.g. [t̀t]. According to Penttilä (1963, p. 21) the syllable boundary of a geminate in Standard Finnish is also situated in its final section so that the greater part of it belongs to the preceding syllable. In his phonetic study on the dialect of South Ostro-bothnia, Laurosela (1922, p. 107–109) also defined the boundaries of geminate components – but on the basis of objective duration. He thinks it probable that the final part of an inter-vocal geminate is pretty much the same in duration as a short (=single) consonant appearing in a corresponding position, but he admits that in the curves of geminates no boundary between the parts can be observed. As an alternative assumption, Laurosela suggests that the duration of the latter component of a geminate might be as long as the duration of a corresponding single consonant which is the latter component in clusters of two different consonant. According to his material both starting points would reach the same results. On the basis of the ratio between geminate duration and the duration of a corresponding single consonant, Laurosela distinguishes between initially long, initially half-long and initially short geminates which he represents thus: k̄k m̄m / k̀k m̀m / k̆k m̆m. These geminate groups, based on comparisons of measured durations, do not correspond to similar quantity degrees defined audiophonetically: Laurosela's initially long, initially half-long and initially short geminates can be subjectively quite 'normal', ordinary geminates. Although the quantity degrees of geminate consonants defined by Laurosela are arbitrary and insignificant as such, the idea of comparing the duration of a geminate consonant to the duration of the corresponding single consonant, or, as he also suggests, to the duration of the latter member of a corresponding cluster of two different consonants, is of practical value. Such comparisons will be made later in chapter 3.7.2.4.

3.7.2.3. The durational distinction between different single and geminate consonants

A comparison of the durational distinctions between consonant segments representing single and double or single and geminate consonants in different structures is not as simple as with vowels because the various single consonants differ from each other considerably more in their intrinsic duration than do vowels. While the relative durational difference between various vowel phonemes in the same position is generally much less than 20 %, the durational difference between two single consonants in the same position can be over 100 % of the shorter duration. When the absolute durational differences between various double consonants are, according to the analysis of nonsense words, about the same as those of single consonants and thus relatively small, even the ratio C/CC calculated from the minimal pairs greatly depends on what part consonants long or short in intrinsic duration have played in the compared word groups. In the structures of the nonsense word series CVCV and CVCCV, the average duration of the intervocal single consonants /ptksmnlr/ is 76msec and the mean of the average durations of the corresponding geminates is 164msec: so a double consonant, according to this, is on average more than twice (2.2) as long as a single consonant. If the average duration of all single and double consonants appearing on the boundary of the first and second syllable is calculated from the list of meaningful sentences, the ratio between double and single consonants is almost exactly 2. The average duration of single consonants is 78msec, double consonants 158msec, therefore C/CC is 2.02. On average one can therefore say that a double consonant is about twice as long in duration as the corresponding single consonant.

The means of the average durations of the consonants in the various word structure pairs of the sentence list and the corresponding C/CC ratios of the various consonant phonemes are presented in Table 10. The phonetic environment is slightly different with different consonants. For example, the r/rr ratio is based on the means of three words (thus $N = 30$) in all of which the first syllable vowel is /i/; in the n/nn comparison, on the other hand, the vowels of all the words are /a/ vowels whereas in other comparisons there are words containing several different vowels.

Because of the irregularity of the phonetic frame there is no reason to test the statistical significance of the ratios. However, the ratios are quite clearly

TABLE 10. The C/CC ratio of different consonant phonemes

segmental phoneme	single $\overline{X}$(msec)	double $\overline{X}$(msec)	lengthening (msec)	C/CC ratio	
/p/	97	194	97	1:2.00	mean ratio 1:1.99
/t/	88	186	98	1:2.11	
/k/	105	194	89	1:1.85	
/s/	81	163	82	1:2.01	
/m/[1]	–	–	–	–	
/n/	47	123	73	1:2.61	mean ratio 1:2.50
/l/	49	118	69	1:2.43	
/r/	42	103	61	1:2.45	

[1] The material of the present work was lacking in minimal pairs with a m/mm distinction.

indicative and correspond to the ratios of single and double consonants in the nonsense words. The consonants fall into two groups based on the C/CC ratio. The obstruents /ptks/ are characterized by relatively small C/CC ratios, while the nasal /n/ and liquids /l r/ commonly have a C/CC ratio which is clearly bigger than the former one. The nasal /m/, which was not represented in a single minimal pair in the material for the purpose of its C/CC calculation, would, according to the nonsense words, come somewhere between those groups. It is interesting to note that the geminate of the consonants which are longest in their intrinsic duration is, on average, twice as long in duration as a single consonant in a corresponding position, but the geminates of consonants which are shorter in their intrinsic duration are, on average, two and a half times the length of the corresponding single consonant. If, in the description of geminate consonants, we also used the length chromene which, in the case of geminates, is combined with the various consonant phonemes, we could imagine that on the production level it has its own corresponding articulation command which tends to be constant in duration. Ideally, the duration of a geminate consonant would be the intrinsic duration of the consonant + the standard 'lengthening' of its duration in which case the durational differences between double consonants could be expected to be the same as between single consonants. However, it is not possible to show the existence of this kind of representative of the standard durational 'lengthening command' of the chroneme on the phonetic production level, which would account for the different C/CC ratios of different consonants. On the other hand, the durational relationships of consonants can be affected by the tendency towards constant duration of some unit larger than an individual phoneme segment. If it is supposed that, for example, a word or measure representing a structure model forms some kind of independent timing unit which has its own characteristic vowel and consonant segment timing pattern, the differences in C/CC ratios accord perfectly with expectations: the doubling of a consonant which is short in its intrinsic duration is not enough to 'fill' the duration of the consonant segment characteristic of the structure, so it must be lengthened more; the doubling of a consonant which is long in duration, on the other hand, can, in fairly quick normal speech tempo, sometimes be too long for the structure's consonant position and the C/CC ratio becomes smaller than the average 1:2. The effects of the durational differences of the various consonants only become clearly visible, however, in relatively quick speech.[1] The C/CC ratio is slightly different in different structure pairs. The change in durational relationships between double and single consonants of the group /p t k s/ in different word structures is presented in the following table:

TABLE 11. The CC/C ratio of /ptks/ in four structure pairs

compared structures	/p/	/t/	/k/	/s/	mean
CVCV/CVCCV	1.77	1.82	1.67	1.87	1.78
CVCVV/CVCCVV	1.98	2.26	1.88	2.05	2.04
CVCVVC/CVCCVVC	2.03	2.21	2.05	2.00	2.06
CVCVCV/CVCCVCV	2.26	2.18	2.22	2.00	2.16

Just as durational differences of different vowels must obviously be considered as automatic consequences of articulatory action of varying complexity and speed, and not deliberately articulated, it can also be supposed that all the

[1] The results of Laurosela's (1921, p. 104–110) and Palomaa's (1943, p. 175–) measurements also seem to indicate that, by slowing down the speech tempo, the durational differences between different single and double consonants are reduced (cf House's observation concerning vowel differences 1961, p. 1175).

different consonants, too, are of equal duration on some 'deeper' level of production, perhaps in production commands of units larger than an individual phoneme. The durational differences of consonants appearing in the acoustic output would also be like the durational differences of different vowels, apart from modifications caused by different articulation times characteristic of different consonants. Various durational compensations take place within a unit larger than a phoneme during the combination of intrinsically different vowels and consonants: long consonants shorten the surrounding vowels and the duration left unused by a consonant which is short in intrinsic duration lengthens the surrounding vowels. This same idea can also be applied to the examination of geminate consonant duration. The first member of a geminate consonant can be seen, in terms of structure, to be in the same position as a vowel preceding a single consonant: just as a single consonant, short in intrinsic duration, lengthens the preceding vowel, the latter member of a geminate, short in intrinsic duration, lengthens the preceding member. Phonetically this is expressed in the following way: a double consonant /nn 11 rr/ is durationally longer than the sum of two single /n/'s etc. in a corresponding phonetic environment. Similarly, the latter component of a geminate plosive shortens the initial part of the plosive just as a single plosive shortens the preceding vowel. This in manifested in such a way that the geminate plosive can be durationally shorter than the sum of the durations of two single plosives.

However, a geminate consonant is phonetically only one long consonant segment in which it is not possible to show the actual durational relationships of the parts of the consonant. On the other hand, it is possible to speculate about the subjective analysis of geminates if one places them alongside clusters of two different consonants in which a phonetic definition of syllable boundary (segment boundary) may be made. If the boundary of the components of the cluster is situated, for example, in the final part of the intervocal consonant segment, it may be supposed, by analogy, that the greater part of the duration of a structurally equivalent geminate belongs, as suggested, in the previous syllable.

3.7.2.4. Consonant clusters

Of clusters of two different consonants the combinations /ts ks st nt ns ŋk/ were chosen, more or less at random, for this study. The average durations of the components of each cluster, measured in the structures CVC_xC_yV,

CVC_xC_yVC, CVC_xC_yVV, $CVVC_xC_yV$, $CVVC_xC_yVC$, $CVVC_xC_yVV$, are presented in the following table. In all the words measured, the vowel preceding and following the cluster was /a/. The ratios of the durations of the members of the consonant clusters is given in percentages.

TABLE 12. Durational distribution in different consonant clusters

consonant cluster	C_x $\bar{X}$ (msec)	C_y $\bar{X}$ (msec)	C_y per cent of C_x	total duration msec	
/nt/	84	61	(73 %)	145	word structures CVCCV CVCCVV CVCCVC
/ns/	73	60	(82 %)	133	
/ŋk/	85	66	(78 %)	151	
/ts/	97	80	(83 %)	177	
/ks/	96	88	(92 %)	184	
/st/	77	93	(120 %)	170	
/st/	52	85	(162 %)	137	word structures CVVCCV CVVCCVC CVVCCVV
/ks/	71	80	(112 %)	151	

Clusters of two different consonants behave within a word in the same manner as geminate consonants. The duration of the whole consonant segment depends on the length of both the preceding and the following vowel segment. Both the duration of geminates and the total duration of clusters is compared to the /CVCCV/ structure, longer before a long vowel segment of the second syllable and shorter after a long vowel segment of the first syllable. The influence of this latter becomes evident if we compare the total durations of the /ks/ and /st/ sequences in the first and second groups. Lengthening of the first vowel segment has shortened the following consonant segment by 19 %, or by approximately the same as the corresponding geminate consonants (cf p. 110). The influence of a preceding vowel is clearly greater on the first consonant of the cluster (shortening it by 27 %) than on the second consonant (shortening it by 9 %). Correspondingly, it will be noticed (p. 113) that a long vowel coming after a cluster lengthens the second consonant of the cluster more than the first.

On the basis of the limited material available, the duration of a consonant cluster would seem to be approximately the same as the sum of the durations of single consonants in the word position corresponding to the members of the cluster. However, the durational ratios between segments seem to vary according to whether there is a short or a long vowel before or after the cluster. It can be seen from the consonant combinations in the table above that a consonant which is short in intrinsic duration, at least /n/, is considerably lengthened in comparison with a corresponding single consonant when it is preceded by a single vowel. At the same time, the second member of the cluster becomes clearly shorter than a corresponding single consonant. The duration of the combination /nt/, for example, is almost the sum (137 ms) of the means of single consonants in a corresponding position, but the durational distribution is not equal to the durations of single consonants: in the CVCV(V) (C) structures the duration of intervocal /n/ is 48 msec, but in the cluster which is in the same phonetical frame it is 84 msec. Correspondingly, the /t/ segment is shorter in duration in the /nt/ combination than an intervocal single /t/ in any structure. In clusters where each member is a phonetically long consonant, a corresponding shift in the ratio of segment duration is not observed: the durations of the segments of the /ts ks st/ combinations correspond approximately to the ratios of the intrinsic durations of the equivalent single consonants. – It could be said that the lengthening of the duration of a phonetically short consonant is caused by the need to make the quantity of the first s y l l a b l e long enough. A phonetically long consonant is, in itself, long enough to produce a long syllable, but duration has to be added to a short consonant. The reason for lengthening could, however, be sought in the demands of word structures without falling back on the syllable concept. A single consonant, a sequence of two consonants, or a sequence of three consonants, can appear between vowels. Only a liquid /l r/ or a nasal /m n ŋ/ can appear as the first member of a CCC sequence, and only plosives and /s/ either as geminates or in some clusters can appear as the latter members. In the consonant segment following a nasal or liquid, there is therefore a structural C/CC distinction, for example *kansa* 'people' /*kanssa* 'with', *kanta* 'base' /*kantta* 'of the cover'. Because the duration of the following vowel segment cannot be used in the CVCCCV structure as an indicator of the biphonemisation of the segment following the liquid and nasal as in the CVCV/CVCCV distinction, the contrast between adjacent consonant segments has developed as an indicator of the $C_{nas,liq}$, $C/C_{nas.liq}$, CC distinction. The first component of a word-medial consonant cluster may lenghthen freely because its duration is not ambivalent in the phonemisation. The distribution of the durations of the intervocal consonant cluster in the (C)VC_xC_yV (V)(C) structures can be considered one of the phonetically characteristic features of Finnish, and one which causes a lot trouble to students of Finnish: if the first segment of the cluster is not lenghthened, the Finnish listener will most often recognize the first vowel of the word as long.[1] Lengthened duration of the first member of the cluster is subjectively perceived as a close contact between vowel and consonant.

As indicated before, the durational ratios between components of the C_xC_y sequence following a single vowel are in contrast in two ways: the lengthened (or in the case of phonetically long consonants, unshortened) duration of the first member indicates that the preceding vowel segment and also the following consonant segment are single phonemes. Because the durational ratios in the consonant segment also influence the phonemisation of adjacent vowel segments, it is understandable that the same ratios are also valid in phoneme sequences where a biphonematic interpretation of the second consonant segment of the cluster is not possible on account of phonotactic restrictions (when the following member of the cluster is a consonant other than (ptks). The durational distribution of a cluster is, moreover, to be considered as a feature of the word structures concerned rather than a sign of the 'phonological quantity' of a certain segment. Once again it may be seen that the duration of sound segments should be treated in Finnish as a phenomenon of word phonology or as a phonetic factor showing syntagmatic composition of phonemes, rather than as representative of the length feature of fixed segmental phonemes in a phoneme sequence or as representative of the mark of the suprasegmental chroneme. (cf Lehiste 1968)

3.7.3. Quantity distinctions and the commutation procedure

An abstract phoneme sequence is to be broken up according to traditional phonemics into individual clusters of distinctive features, or segmental phonemes. On the other hand the acoustic speech output is not to be broken up in

[1] For pedagogical uses of phonetic transcription, the first segment of the cluster could be marked half-long: for example, [kulˑta:] 'of gold' but [kul:ta:] 'it shines'.

the same way into individual fragments of consecutive sound waves. The phonetic cues indicating the phonemic status of a certain sound segment can appear over a larger area than that segment: the formant transitions indicating the point of articulation of consonants are located in the area of the nearby vowel and, for example, in English the important cue indicating voicedness of a word-final consonant is in the duration of the preceding vowel segment and so on. A good example of a phonetic area larger than one sound segment in a distinction considered phonemically as segmental may be found in the manifestations of the distinction long vowel/short vowel (tense/lax vowel), for example in Swedish, Norwegian and German. Both in the Scandinavian languages and in German the phonetic area of distinction is the sequence of vowel and following consonant: when the vowel is long and the following consonant short, the vowel is recognized as /long/. If we perform a so-called commutation, the changing of a long vowel to a short vowel is automatically accompanied by the lengthening of the following consonant, for example Swedish [fi:na/fin:a].[1] Estonian is also a favourable subject for the examination of the phonetic domain of length distinction, even though its character differs noticeably from the previous example. The length distinction of Estonian vowels is said to be relevant only in the first, stressed syllable of a word. The vowel of the second syllable is phonetically short, if the vowel of the first syllable has the III degree of length; in the I and II degrees of length the vowel of the second syllable is described as being phonetically half-long (cf. note 1 on page 22). Regardless of the manner in which the three distinctive durations of Estonian are treated on the phoneme level, it is an undisputed fact, that the quantity degrees of Estonian need a two syllable sequence as their phonetic domain.[2] A description of Estonian quantity patterns is necessarily incomplete if restricted to segmental phonemes or even syllables (Lehiste 1960, p. 63). Ravila (1961) verifies from the quantity system of Estonian (and Lappish) that different degrees of length function in many cases just like indicators of certain structure types. Lehiste (1965, p. 455–456) has also paid attention to the function of quantity as a higher-level phonological unit in Estonian: »In the phonological hierarchy of Estonian there are two intermediate levels between the phoneme and the

[1] For the intercorrelation between vocal and consonant length in Swedish: see Elert 1965; in Norwegian Fintoft 1961; in Icelandic Benediktson 1963.

[2] From the point of view of perception of the quantity of a word the ratio of the duration of the first syllable to the second syllable vowel is the most significant feature, too (Liiv 1962, p. 286).

phonological word: the syllable and the dissyllabic sequence. Quantity is the factor which relates the units within the hierarchy to each other.»

The phonetical manifestations of the three quantity degrees of Estonian are undoubtedly fairly complex. For example the words [lina/lin: a/lin:ˑ a] or [saDa/sa:Da/sa:ˑDa] are not phonetically distinguished from each other only by the different duration of the intervocal [n] or of the [a] of the first syllable but by the complexity of many different phonetic features. For this reason one duration degree cannot be substituted for some other on the phonetic level just by changing the phonetic duration of the segment in question (Ravila 1961, p. 349). Compared to phonemisation solutions for the quantities of Estonian, the Finnish quantity sytem seems to be simple: the long quantity of both vowels and consonants can in phonemisation simply be seen as a sequence of two similar segmental phonemes. A phonemically simple solution, however, does not mean that the phonetic durations of sound segments are in one-to-one correlation to the corresponding points in the phoneme sequence. Ravila suggests (1961, p. 347) that in Finnish one could »perform substitution between the long and short degree of just one phoneme whilst the remaining environment stays the same. In the minimal pairs, which are numerous, the quantity of an individual phoneme is, therefore, a distinctive factor, and we do not require syllable, speech measure or any other entity greater than a phoneme as the basic unit of quantity differences.»

The foregoing seems to mean, among other things, that it would be possible to change the duration of one sound segment without interfering at all with the phonetic frame or with the durations of the other sounds of a word. This is not the case in Estonian, where a change in the phonemic quantity of one sound automatically causes related changes in duration of other sounds, in the intonation etc.

The following comparisons of average segmental durations of word structures and the perception tests of chapter 4 indicate that in some cases replacing a double sound by a single sound in a certain sound position of a word is possible in Ravila's sense. But it is obvious that in the substitution pairs presented by Ravila *tuli* 'fire'/*tulli* 'customs' and *tuli*/*tuuli* 'wind' the change is not possible at all without a change in environment or without a change in the intended perception of the word. Comparison of word structures will also show that the phonetical duration of the sound segments of Finnish words cannot be treated thoroughly without reference to frames larger than one sound segment.

3.7.4. Comparison of segmental durations in minimally distinctive word structure pairs

The groups of compared minimal word pairs have been chosen in each comparison group so that the intrinsic duration of sounds and the influence of neighbouring phonemes on each other's duration can be disregarded when examining the durational differences of various word structures. In each comparison only word pairs have been chosen for the comparison groups in which the composition of different segmental phonemes is the same, in other words, minimal pairs differing from each other only in the phonemic function of one qualitative sound segment. For example in the comparison of the durational patterns of the structures CVCV and CVCVV the durational averages of all the measured allophones of the phonetic C_1 V_1 C_2 and V_2 segments are calculated from the words *tapa* 'manner', *lika* 'dirt', *ripi* 'birch', *sata* 'a hundred' etc. which have been compared to the means and distributions of duration of the common allophones of the words *tapaa* 'he meets', *likaa* 'of dirt', *ripii* 'it is sprinkling', *sataa* 'it is raining' etc. Then the influence of the intrinsic duration of different sound combinations etc. on the means of each group is similar and observed durational differences can be supposed to be caused by the differences in word structures.

The durations of the sound segments in the same position of separate comparison pairs are not, however, quite comparable, because in different comparison series different segmental sounds can be stressed in different ways in some positions: in one series a C-position may have a greater number of plosives, in another there may be, for example, more nasals so that the means can differ considerably, even though the segments in question are of the same durational value in the structures. The comparison results of the durations of phonetical sound segments within minimally differing structures are shown in tables $13._{1-12}$. The differences in the durational patterns of the structures are also presented graphically in connection with the discussion of the individual structure pairs.

The duration of each sound segment of the structures under comparison is shown in milliseconds in the tables along with the number of individual cases (N), in which the mean is calculated. Ideally, the number of cases should be the same in all the segments of each group; those small differences in the number of cases appearing in some comparison which hardly influence the resultant relationships have several causes: one informant has misread a word, or in some cases one word has not been measured because of disturbance during the recording etc. The number of so-called omissions in the whole material, however, is very small.

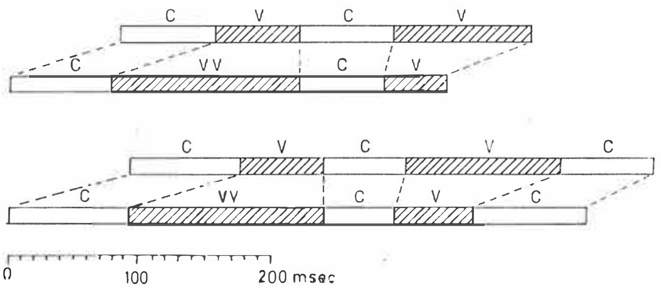
The differences in the means of the sound segment durations is represented in milliseconds and the statistical significance of the difference by asterisks, or by a percentage mark when the difference is not significant, but sufficiently plausible so that it can be taken as indicative. One should notice that 'significance' does not refer to the absolute amount of differences in milliseconds, but to what extent it is probable that the difference is of the same tendency in any new case in the groups of compared variables. Secondly, it should be remembered, that statistical significance has not necessarily anything to do with perceptual significance. A difference which is statistically strongly significant can have no significance at all in perception. In principle, of course, a difference which is not significant from a statistical point of view can be linguistically relevant, but, on the other hand, it is to be presumed that a difference which is not repeated regularly in a series of structurally identical oppositions could not be used as an identity cue on the perception level either.

1. CVCV/CVVCV ; CVCVC/CVVCVC

for example:

sama 'the same'/*saama* 'received'; *saman* 'of the same'/*saaman* 'of the received'

Figures 6_{1-10} (pp. ??–??). Comparisons between the average segmental durations in various minimally differing word structure pairs. The differences are discussed in the corresponding chapters of the text and the statistical significance of the differences is tested in tables 13_{1-10}.



Comparisons of the durations of the sound segments for the *sama/saama* and *saman/saaman* type minimal pairs are shown in the table $13._1$ and in the figure above. In both compared groups the initial consonant preceding a long vowel is durationally slightly longer on average, but the difference is not at all significant in the first one and even in the second one it is only on the 20% level. The reason for the smaller number of measured cases in the C_1-segment is that some of the readers have assimilated the word-boundary consonant combination /nl nr nv nm/ as /ll rr vv mm/. This does not permit the measurement of the duration of a word-initial single consonant. This kind of assimilation is accidental with all the speakers with the exception of the assimilation /nm>mm/, which is an acceptable feature of the standard language. Beside word-boundary assimilation omissions, the sound environment of different sentence frames also influences the comparability of the durations of word-initial consonants: in some comparisons the durational difference of initial consonants may be caused by a different word-boundary consonant cluster in the frame sentences of the word pairs to be compared; for example, a word-initial plosive following a word-final /n/ is noticeably shorter in duration than an intervocal /t/ appearing on the word boundary. Because of these factors the durational differences of initial consonants of a word should not be immediately taken as significant, even though the t-test would show that the difference is statistically significant.

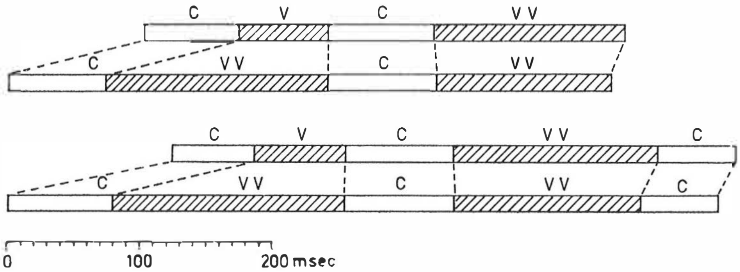
With the exclusion of the initial consonants the durations of the structures of all the sound segments differ significantly from each other. The smallest difference is in the C_2-segment, which is significantly longer in the first group ($p < 1\%$) and non-significantly longer ($p < 10\%$) in the latter group after a single than after a double vowel. The difference in both cases is only 8msec. The most noticeable durational differences in the structures are in the V_1 and V_2 segments: when, in place of a single vowel, a double vowel appears in the V_1 position in the structure, the duration of both V_1 and V_2 segments also changes in about the same proportion: the first one is durationally doubly lengthened and the second one is durationally shortened by half. The V/VV ratio in the compared group is about 44 % (in the first 44 % and in the second 43 %) and the shortening of the V_2 is on average 50 % (52 % and 49 %).

Phonemic doubling of the first vowel segment lengthened the whole CVVCV-word by only 20 msec compared to the CVCV word and the CVVCVC-word by 41 msec compared to the CVCVC word. The CVCV(C)/CVVCV(C) distinction represents in durational distribution above all a shift in duration from the vowel of the second syllable to the vowel of the first syllable. The substitution /s-a-ma/s-aa-ma/, which on the phonemic level means an addition of one segmental phoneme (or of a long chroneme), does not mean, however, on the phonetical level, the doubling or mere increase in the duration of the corresponding sound segment, *ceteris paribus*, in the phonetical sense, but substantial change in the timing of the sound segments of whole structure.

2. CVCVV/CVVCVV ; CVCVVC/CVVCVVC

for example:

samaa'of the same' /*saamaa* 'received'; *samaan* 'to the same'/ *saamaan* 'to receive'
(3. inf.part.) (3. inf. ill.)



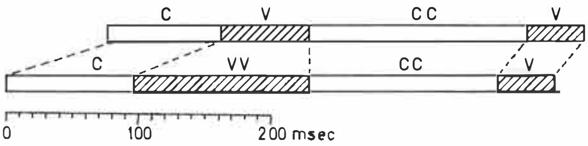
The durational differences in words containing a single and double first syllable vowel, in which the second vowel is long, appear in table 13_2 and in the figure above. In the *samaa/saamaa* and *samaan/saamaan* types of distinction, the duration of the intervocal single consonant does not vary. It is 2msec longer in both series after a double vowel, but the difference in the means is quite accidental. The duration of a word-initial consonant would, according to the test, remain unchanged in the CVCVV/CVVCVV-distinction, but would be significantly smaller in the CVVCVVC-structure than in the CVCVVC-structure. This unexpected result may, however, be due to some geminate consonants at the word-boundary having, in a few cases, shifted during the calculation of the duration of the previous C_1. There are no differences in the duration of the final consonants of structures ending in consonants.

The most noticeable difference in the structures under consideration is, as expected, in the duration of the V_1-segment. In both comparison groups the double vowel is over 2.5 times as long in phonetic duration as the single vowel. The durational differences between the second-syllable double vowels are relatively small: after a first-syllable double vowel the second-syllable double vowel is about 10 % shorter in duration than the corresponding vowel after a single V_1. The phonemically double vowel is not as long in absolute duration in the second syllable as in the first. In the CVVCVV(C)-structures under consideration the duration of the second-syllable double vowel is about 80% of the duration of the first-syllable double vowel.

3. CVCCV/CVVCCV

for example:

takka 'fire-place'/*taakka* 'burden'

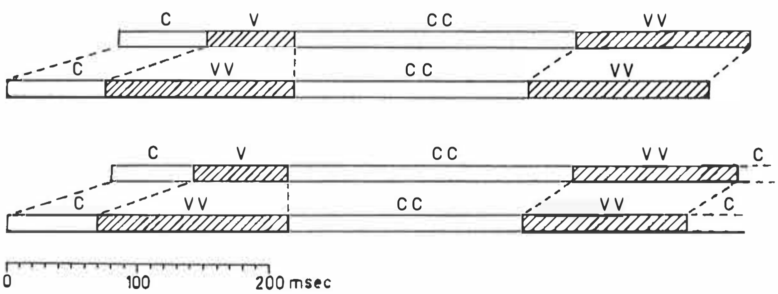


The number of perfect minimal pairs suitable for the CVCCV/CVVCCV comparison is small, but the groups can be considered representative, because their average durations and durational patterns correspond very well to the average durations of the corresponding structures shown in table 14, which are calculated on the basis of 266 and 80 measured cases. The duration of the initial consonant of this structure comparison is 10msec longer before a long vowel, but the difference is not statistically significant ($p < 20\ \%$). The duration of the second syllable vowel is the same in both structures. The double vowel of the first syllable is exactly twice as long in its duration as the single vowel in the equivalent position. The duration of the intervocal double consonant is 13 % shorter after the double vowel of the first syllable than after the first- syllable single vowel.

4. CVCCVV/CVVCCV ; CVCCVVC/CVVCCVV

for example:

takkaa 'of the fire-place' /*taakkaa* 'of the burden' ; *takkaan* 'into the fire-place'/ *taakkaan* 'into the burden'

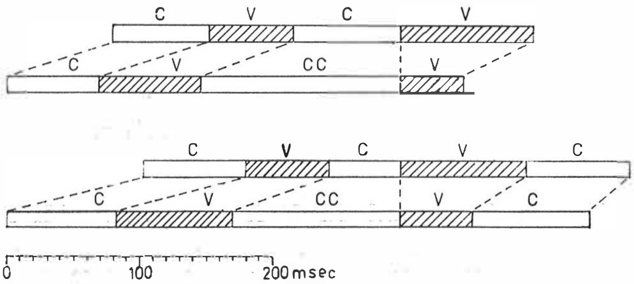


In both comparison series the word-intitial consonant preceding the double vowel is longer than that preceding the single vowel but the difference is not statistically significant in either of the series. However, the initial consonant is found to be longer in duration in all the structure pairs when it precedes a double than when it it precedes a single vowel, even though the difference is not statistically significant in most cases. In the CVCCVV/CVVCCVV-series the difference is almost significant ($p < 5\%$) ; the difference in the means is in both series under consideration slightly under 10msec in favour of the double vowel structure.

There are no significant differences in the durations of the second syllable vowel in either of the comparison series. In the durations of the final consonant in the second comparison series there are, moreover, no significant differences. In both series the double vowel of the first syllable is slightly more than twice the length of the single vowel in its duration. The V/VV ratio is in the first series 46 % and in the second 50 %. The intervocal double consonant is significantly shorter in its duration after a double vowel than after a single vowel. The durational reduction is 18 % in both series.

The double vowel is durationally shorter in the second syllable than in the first syllable. In the CVVCCVV and the CVVCCVVC-structures, the double vowel of the second syllable is in the first series durationally 96 %, in the second 86 %, of the duration of the double vowel of the first syllable.

5. CVCV/CVCCV ; CVCVC/CVCCVC
for example:
napa 'navel' /*nappa* 'leather' ; *pysyt* 'you stay' / *pyssyt* 'guns'



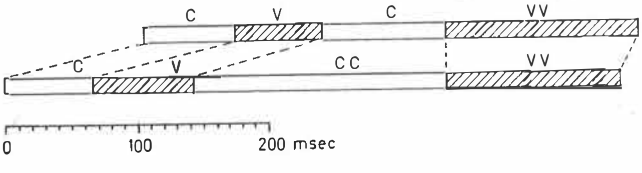
In the structure pairs under comparison the initial consonant is longer in duration in the geminate consonant member CVCCVC but shorter in the structure CVCCV. In neither of the pairs is the difference in duration of the initial consonant statistically significant. On the other hand, the durational differences in all the other sound segments are statistically strongly significant. In the pair CVCV/CVCCV the first syllable vowel preceding the geminate consonant is 20 %, and in the pair CVCVC/CVCCVC 40 % longer than when preceding a single consonant. The intervocal geminate consonant is about 1.9 times in the first comparison series and in the second 2.2 times the duration of the single consonant. The different durational ratio between the single and double consonants as well as the considerably different absolute duration in the comparison series CVCV/CVCCV and CVCVC/CVCCVC is caused by the fact that in the first series more consonants of the C_2-position were obstruents (ptks), whose intrinsic duration is greater and C/CC ratio smaller than those of the other consonants. After the geminate consonant the single vowel of the second syllable is in duration on average only half the duration of the vowel which comes after the single consonant. The reduction is 54 % in the first pair and 44 % in the second. In the second pair the word-final consonant appearing after the geminate consonant is longer in duration than that appearing after a single consonant. In each geminate structure the single vowel of the second syllable is shorter than the single vowel of the first syllable: in duration V_2 is 62 % in the CVCCV structure and 61 % in the CVCCVV-structure of the duration of the V_1.

In the CVCV(C)/CVCCV(C) -distinction as in the CVCV(C) /CVVCV(C)-distinction, the shorter or longer duration of a given sound segment does not merely correspond to the phonemic opposition of the single phoneme to the double phoneme. Both in the first syllable V/VV distinction and in the intervocal consonant C/CC distinction the change of duration in the phonemically single second syllable vowel is relatively of almost the same size as the change in duration of the sound segments representing single and double phonemes.

6. CVCVV/CVCCVV

for example:

takaa 'behind' /*takkaa* 'of the fire-place'

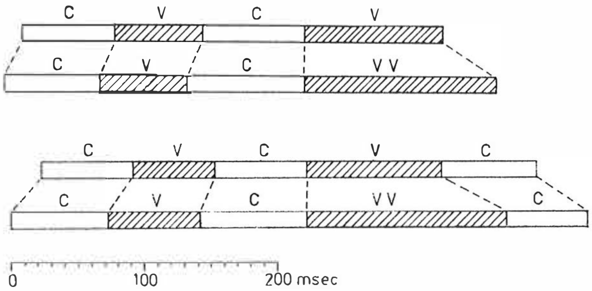


The durational difference between the initial consonants is not significant in the distinction CVCVV/CVCCVV. The first syllable vowel is 18 % longer in the geminate consonant structure than in the single consonant structure and the second syllable vowel is 8 % shorter in the geminate consonant structure than in a single consonant structure. The duration of the intervocal single consonant is in the ratio 1:2.3 to the duration of the geminate consonant.

7. CVCV/CVCVV ; CVCVC/CVCVVC

for example:

sata 'hundred' / *sataa* 'it is raining'; *takan* 'the fire-place's' / *takaan* 'I guarantee'

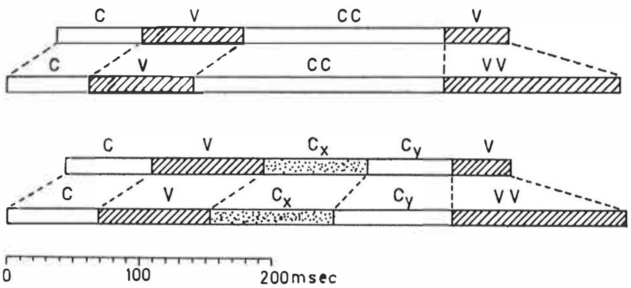


Apart from the final consonant of the CVCVVC-structure all the sound segments are durationally longer in structures where the second syllable vowel is phonemically double. The single vowels of the first syllable are, however, of the same duration in the CVCV and CVCVV structure. The intervocal single consonant in both comparison series is 16 % longer before the double vowel than before a single vowel. The durational ratio of the second syllable single vowel to the second syllable double vowel is 1 : 1.38 in the CVCV/CVCVV distinction and 1 : 1.48 in the CVCVC/CVCVVC distinction.

8. CVCCV/CVCCVV ; CVC_xC_yV/CVC_xC_yVV

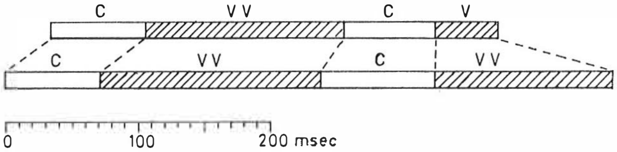
for example:

takka 'fire-place' / *takkaa* 'of the fire-place' (part.); *taksa* 'rate'/*taksaa* 'of the rate' (part.)



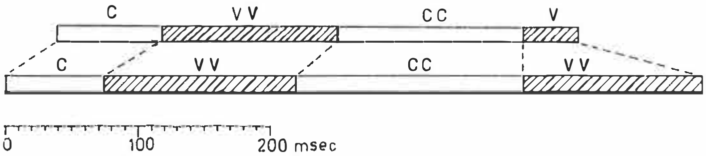
The durational differences between initial consonants and first syllable single vowels are not significant. On the other hand the intervocal geminate consonant and the intervocal consonant cluster lengthen before the double vowel, the former by 25 % and the latter by 28 % . The first consonant in the cluster lengthens slightly less than the second. In both comparison series the double vowel of the second syllable is almost three times as long as the single vowel of the second syllable.

9. CVVCV/CVVCVV
   for example:
   laaka 'flat' /*laakaa* 'of flat'



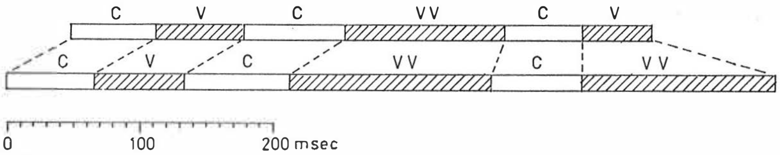
Both the double vowel of the first syllable and the intervocal single consonant are durationally longer before the second syllable double vowel than before the second syllable single vowel. The increase in duration of V_1 is 12 % and the increase of C_2 29 %. The duration of the second syllable single vowel is in the ratio 1:2.81 to the duration of the corresponding double vowel. In the CVVCVV-structure the duration of the double vowel of the second syllable is 80 % of the duration of the first syllable vowel.

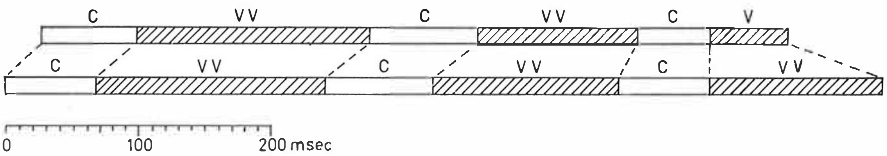
10. CVVCCV/CVVCCVV
    for example:
    taakka 'burden' / *taakkaa* 'of a burden'



All but the initial consonant are durationally longer in the CVVCCVV than in the CVVCCV -words. The increase of 10 % in the duration of the double vowel of the first syllable is, however, statistically only almost significant. The lengthening of the intervocal geminate consonant preceding the double vowel is 23 %. The double vowel of the second syllable is durationally over three times as long as the corresponding single vowel.

11. CVCVVCV/CVCVVCVV : CVVCVVCV/CVVCVVCVV
 for example:
 takaama 'guaranteed' / *takaamaa* 'of guaranteed': *vaakaama* /'balanced' *vaakaamaa* 'of balanced'



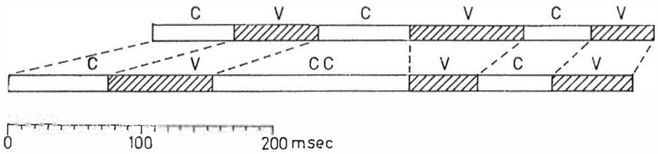


In these two compared structure pairs, in which there is a single vowel/double vowel distinction in the third syllable, there are no significant differences in the duration of the first three sound segments in either case. However, in both cases the durations of both the intervocal single consonant and the double vowels of the penultimate syllable are quite significantly longer before the double vowel of the last syllable than before the single vowel. The increase in the V_2-segment is 26 % in the CVCVVCVV structure, 17 % in the CVVCVVCVV structure, the increase in the duration of the intervocal consonant in the former is 17 %, and in the latter 24 %. The durational ratio of the single vowel to the double vowel of the first syllable in the CVCVVCV/CVCVVCVV distinction is 1:2.8 and in the CVVCVVCV/CVVCVVCVV distinction 1:2.2. The duration of the double vowel of the third syllable in the CVCVVCVV structure is 96 % of the duration of the double vowel of the second syllable. In the CVVCVCCVV-structure the corresponding percentage is 94 %. In the CVVCVVCV-structure the duration of the double vowel of the second syllable is 68 % of the duration of the double vowel of the first syllable, in the CVVCVVCVV structure it is 81 %.

12. CVCVCV/CVCCVCV

for example:

takana 'behind'/*takkana* 'as a fire-place'



The pair CVCVCV/CVCCVCV, in which there is a single and double consonant distinction on the boundary of the first and second syllable, can be compared to the corresponding distinction of the two bisyllabic structures dealt with in chapter 3.7.4.5. Significant changes in the average durations appear in both, besides the change in the consonant segment itself, in the duration of the single vowel preceding it and following it. In the CVCVCV/CVCCVCV distinction the duration of the vowel in the third syllable also changes significantly. The difference of durations between the initial consonants of the structures in these comparison pairs may be partly caused by different consonant combinations in the sentence frames.

Both V_1 and V_3 are longer in the geminate consonantal structure than in the CVCVCV-structure: the increase in duration of the former is 26 %, and of the latter 30 %, whereas V_2 is shorter after the geminate than after the single consonant. The decrease is about 29 %. The phonemically double C_2 is about 2.2 times as long in duration as the corresponding single consonant.

A feature worth noticing in the CVCV/CVCCV distinction concerned changes in phonetic duration contrast in the first and second syllables, which are phonemically of the same value: in the CVCV-structure the V_2 is longer than the V_1 in duration, in the CVCCV-structure the contrast ratio is the reverse. The same kind of difference in contrast ratio in the single vowels is also observed in the CVCVCV/CVCCVCV-pair, where the vowel of the first syllable also plays a part in the change: in the CVCVCV-structure $V_1 < V_2 > V_3$, in the CVCCVCV structure $V_1 > V_2 < V_3$.

TABLE13$_{1-12}$ Comparisons of segmental durations in various minimally differing word structure pairs

1.a) Compared structures CVCV – CVVCV

Sound segment	Structure CVCV Mean (msec)	n	Structure CVVCV Mean (msec)	n	Difference of means (msec)
C_1	72	84	76	90	–4^{non}
V_1	63	109	143	102	–80****
C_2	72	109	64	102	+8***
V_2	104	109	48	102	+56****

1.b) Compared structures CVCVC – CVVCVC

Sound segment	Structure CVCVC Mean (msec)	n	Structure CVVCVC Mean (msec)	n	Difference of means (msec)
C_1	84	48	90	51	–6^{non} ($p<20\%$)
V_1	63	58	148	59	–85****
C_2	62	58	54	59	+8^{non} ($p<10\%$)
V_2	118	58	60	59	+58****
C_3	70	58	86	59	–16****

2.a) Compared structures CVCVV – CVVCVV

Sound segment	Structure CVCVV Mean (msec)	n	Structure CVVCVV Mean (msec)	n	Difference of means (msec)
C_1	72	99	73	109	–1^{non}
V_1	67	103	168	109	–101****
C_2	81	103	83	109	–2
V_2	145	103	132	109	+13****

2.b) Compared structures CVCVVC – CVVCVVC

Sound segment	Structure CVCVVC Mean (msec)	n	Structure CVVCVVC Mean (msec)	n	Difference of means (msec)
C_1	63	75	79	76	–16****
V_1	68	78	176	80	–108****
C_2	82	78	84	80	-2^{non}
V_2	155	78	140	80	+15****
C_3	58	78	58	80	0

3) Compared structures CVCCV – CVVCCV

Sound segment	Structure CVCCV Mean (msec)	n	Structure CVVCCV Mean (msec)	n	Difference of means (msec)
C_1	86	29	96	50	-10^{non}
V_1	66	29	132	30	–66****
C_2	165	29	143	30	+22****
V_2	43	29	42	30	$+1^{non}$

4. a) Compared structures CVCCVV – CVVCCVV

Sound segment	Structure CVCCVV Mean (msec)	n	Structure CVVCCVV Mean (msec)	n	Difference of means (msec)
C_1	67	59	76	58	–9*
V_1	65	59	143	58	–78****
C_2	216	59	178	58	+38****
V_2	132	59	137	58	-5^{non}

4.b) Compared structures CVCCVVC – CVVCCVVC

Sound segment	Structure CVCCVVC Mean (msec)	n	Structure CVVCCVVC Mean (msec)	n	Difference of means (msec)
C_1	61	27	69	44	-8^{non}
V_1	72	46	145	50	-73****
C_2	216	46	178	50	+38****
V_2	125	46	125	50	0
C_3	64	39	61	40	$+3^{non}$

5. a) Compared structures CVCV – CVCCV

Sound segment	Structure CVCV Mean (msec)	n	Structure CVCCV Mean (msec)	n	Difference of means (msec)
C_1	73	105	69	137	$+4^{non}$
V_1	64	138	77	139	-13****
C_2	81	138	152	139	-71****
V_2	105	138	48	139	+57****

5. b) Compared structures CVCVC – CVCCVC

Sound segment	Structure CVCVC Mean (msec)	n	Structure CVCCVC Mean (msec)	n	Difference of means (msec)
C_1	77	96	82	108	-5^{non} ($p \lessapprox 10\%$)
V_1	62	105	87	110	-25****
C_2	57	105	128	110	-71****
V_2	95	105	53	110	+42****
C_3	79	87	89	83	-10****

6) Compared structures CVCVV – CVCCVV

Sound segment	Structure CVCVV Mean (msec)	n	Structure CVCCVV Mean (msec)	n	Difference of means (msec)
C_1	68	164	66	199	$+2^{non}$
V_1	65	196	77	199	–12****
C_2	93	196	192	199	–99****
V_2	145	196	132	199	+12****

7.a) Compared structures CVCV – CVCVV

Sound segment	Structure CVCV Mean (msec)	n	Structure CVCVV Mean (msec)	n	Difference of means (msec)
C_1	69	200	71	201	-2^{non}
V_1	66	247	66	244	0
C_2	77	247	89	244	–12****
V_2	105	247	145	244	–40****

7. b) Compared structures CVCVC – CVCVVC

Sound segment	Structure CVCVC Mean (msec)	n	Structure CVCVVC Mean (msec)	n	Difference of means (msec)
C_1	68	180	73	172	–5*
V_1	62	196	69	197	–7****
C_2	69	196	80	197	–11****
V_2	102	196	151	197	–49****
C_3	71	166	61	175	+10****

8. a) Compared structures CVCCV – CVCCVV

Sound segment	Structure CVCCV Mean (msec)	n	Structure CVCCVV Mean (msec)	n	Difference of means (msec)
C_1	64	135	61	138	$+3^{non}$
V_1	77	136	78	138	-1^{non}
C_2	153	136	191	138	–38****
V_2	48	136	134	138	–86****

8. b) Compared structures CVCxCyV – CVCxCyVV

Sound segment	Structure CVCxCyV Mean (msec)	n	Structure CVCxCyVV Mean (msec)	n	Difference of means (msec)
C_1	65	60	69	60	-4^{non}
V_1	85	60	85	60	0
C_2	143	60	183	60	–40****
V_2	45	60	131	60	–86****
C_{2x}	78	60	93	60	–15*
C_{2y}	64	60	91	60	–27****

9. Compared structures CVVCV – CVVCVV

Sound segment	Structure CVVCV Mean (msec)	n	Structure CVVCVV Mean (msec)	n	Difference of means (msec)
C_1	71	92	73	99	-2^{non}
V_1	150	93	168	100	–18****
C_2	69	93	89	100	–20****
V_2	48	93	135	100	–87****

10. Compared structures CVVCCV – CVVCCVV

Sound segment	Structure CVVCCV Mean (msec)	n	Structure CVVCCVV Mean (msec)	n	Difference of means (msec)
C_1	80	30	74	38	+6non
V_1	134	40	147	38	–13*
C_2	142	40	174	38	–32****
V_2	43	40	137	38	–94****

11. a) Compared structures CVCVVCV – CVCVVCVV

Sound segment	Structure CVCVVCV Mean (msec)	n	Structure CVCVVCVV Mean (msec)	n	Difference of means (msec)
C_1	63	40	65	39	–2non
V_1	66	49	68	49	–2non
C_2	76	49	80	49	–4non
V_2	121	49	153	49	-32****
C_3	58	49	68	49	–10****
V_3	53	49	147	49	–94****

11. b) Compared structures CVVCVVCV – CVVCVVCVV

Sound segment	Structure CVVCVVCV Mean (msec)	n	Structure CVVCVVCVV Mean (msec)	n	Difference of means (msec)
C_1	72	30	69	28	+3non
V_1	176	30	174	29	+2non
C_2	81	30	82	29	–1non
V_2	120	30	140	29	–20****
C_3	55	30	68	29	–13****
V_3	61	30	132	29	–71****

12. Compared structures CVCVCV – CVCCVCV

Sound segment	Structure CVCVCV Mean (msec)	n	Structure CVCCVCV Mean (msec)	n	Difference of means (msec)
C_1	62	51	77	59	–15****
V_1	63	59	79	60	–16****
C_2	68	59	149	60	–81****
V_2	86	59	51	60	+25****
C_3	51	59	56	60	-5^{non} ($p \leq 20\%$)
V_3	47	59	61	60	–14****

3.7.4.1. Synopsis of comparisons

The comparison series of minimal pairs shows that word structures which differ from each other phonemically in the length of a sound segment in a certain position also differ on the phonetical level in the duration of sound segments other than those representing phonemically single or double sounds. These accompanying changes are in various distinctions of very different sizes both absolutely and relatively. The lengthening of a sound segment in a certain position (by phonemical doubling) seems, however, to have a regular influence on the sound segment in the environment.

3.7.4.2. The influence on the environment of the length of the first syllable vowel

The difference between the durations of double consonants after a single and double vowel seems to be regular. In all the three comparisons a geminate consonant is significantly shorter after a double vowel than after a single vowel. The average shortening is, however, only 16 % of the duration of a geminate consonant following a single vowel. The influence of vowel doubling on the following single consonant is not, however, consistent. If a single consonant after the first vowel is followed by a single vowel in the second syllable, according to the means, the consonant is about 12 % (8 msec) shorter after a double vowel than after a single first syllable vowel. The shortening, however, is not statistically significant in the CVCVC/CVVCVC comparison ($p < 10\%$) and in the CVCV/CVVCV comparison it is significant only on the 1 % risk level. If a single consonant on the syllable boundary is followed by a double vowel, the vowel preceding the consonant has no influence on the duration of the consonant at all.

In all structures where a single consonant appears between vowels the length of the first syllable vowel has a significant influence on the vowel of the second syllable. A second syllable double vowel following a first syllable double vowel is on average 9 % (14 msec) shorter than following a first syllable single vowel. When a single vowel comes after a single consonant as the second syllable vowel, the differences between the vowels following the single or double vowel of the first syllable are considerable: a 2nd syllable single vowel after a double vowel is on average 52 % (57 msec) shorter in duration than after a single vowel of the first syllable.

Therefore lengthening of the first syllable vowel has a shortening effect on the duration of the following sound segment. If the next consonant segment is phonemically long, that is a double consonant, this will become shorter, but if the next consonant is a single consonant, the shortening falls on the following vowel segment.

3.7.4.3. The influence of the length of a second syllable vowel on the environment.

In all the comparisons a double second syllable vowel has lengthened the preceding consonant. The case is exactly the same in the only three syllable structure comparison, which can be dealt with here in connection with two syllable structures. In a part of the comparisons the vowel preceding the consonant has also been significantly lengthened before the double vowel of the second syllable, but this is not a consistent tendency. The absolute lengthening seems to to be slightly bigger in the double consonant than in the single consonant. The percentage of the lengthening of a consonant preceding a double vowel compared to the duration of a consonant preceding a single vowel is on average 22 %. The structure pairs CVCVCV/CVCVVCV of the sentence list are not completely identical in the composition of their phonemes, but one can, however, observe the same tendency in them: a consonant preceding a double vowel is slightly longer (on average 10msec, 14 %) than a consonant preceding a single vowel. Some observations have also been made concerning the lengthening influence of a durationally long vowel on the preceding consonant in Lappish (Äimä 1914, p. 203), in Estonian dialects (Artiste 1941, p. 6) and in the South Ostro-Botnian dialect of Finnish (Laurosela 1922, p. 231).

3.7.4.4. The influence of an intervocal consonant on the environment

In all four compared structure pairs in which words differ from each other in length of intervocal consonant, gemination has a statistically strongly significant influence on both the duration of the preceding vowel, and the subsequent vowel. In all of the cases the vowel preceding a geminate consonant is longer than the vowel before a single consonant. The average lengthening is 16 msec which means, in percentage, 26 %.[1]

A vowel following a geminate consonant is shorter in all the compared cases than a vowel following a single consonant. The shortening is strongly noticeable in structures where both the first and second syllable vowels are single vowels. In these the average shortening is 41 %, whereas the shortening of a second syllable double vowel in the CVCCVV-structure compared to the corresponding duration of the CVCVV-structure is only 8 % .There was not an adequate number of CVVCV/CVVCCV or CVVCVV(C)/CVVCCVVCC comparison pairs included in the material in which the phonemic quantity of the intervocal consonant was the only distinctive factor, therefore the difference between these structures could not be calculated by means of the t-test. The means of the two compared word pairs (*muuta*/ *muutta*, *kiiti* '(he) hurried' /*kiitti* '(he) thanked') of the CVVCV/CVVCCV structures give, however, some indication of the durational distribution: in both the vowel preceding a geminate is longer than the vowel preceding a single consonant: consonant doubling, however, does not seem to have a consistent influence on the duration of the subsequent single vowel. The influence of a geminate consonant between double vowels (in structures CVVCVV/CVVCCVV and CVVCVVC/CVVCCVVC) seems to be the reverse of that in the structures dealt with earlier, even though absolute results cannot be presented because of the small numbers of complete

[1] Sovijärvi has, in his analysis of the schwa-vowels of certain Finnish dialects (1937, p. 17), made observations of the influence of the schwa which seems to reflect the general influence of single and double consonants on the duration of the preceding vowel discussed here: the growth of the schwa which changes the original structure (C)VCCV– – to (C)VCVCV– – (e.g./jalka/'foot' to/jalaka/) has a shortening effect on the duration of the single first syllable vowel.

minimum pairs in the study material. The lengthening of consonant segment between double vowels seems to shorten the preceding durationally long vowel, whereas in the influence of gemination on the duration of the subsequent double vowel no clear tendency can be seen.

3.7.5. A general review of bisyllabic structures

The average distribution of the duration of vowel and consonant segments in bisyllabic word structures in Finnish is graphically presented in figure 7. The respective numerical values with the standard deviations of measured items are presented in table 14. In the graphs the white parts represent the average duration of a consonant segment and the black parts that of a vowel segment in the words of the sentence list which represent each structure. Since the segmental phonemic composition of the test words and the number of measured cases vary in different structures, there is no reason to draw conclusions from durational differences between certain minimal structure pairs on the basis of the small durational differences between segments in the graphs.

On the basis of the total means a number of preliminary observations can, however, be made concerning the general features characteristic of timing patterns of sound sequences in Finnish word structures. The bisyllabic combinations possible in Finnish are placed in four groups in the figure: bisyllabic structures are in the first group, in which a single vowel is the first vowel segment and a single consonant the intervocal consonant. The bisyllabic structures in the second group are those in which a single vowel is followed by a double consonant, in the third group are the structures in which a double vowel is followed by a single consonant and in the fourth group there are the structures in which both the first vowel segment and the following consonant segment represent a double phoneme.

The duration of the word-initial consonant (C_1) does not seem to depend on the following structure. The small durational differences which may be seen between the durations of initial consonants can be incidental and may be caused either by the varying number of intrinsically short and long initial consonants in separate groups or by the fact that the influence of the frame sentence on the duration of the word boundary consonant varies slightly from case to case. On the other hand the duration of final consonants of words obviously depends on the kind of structure the final consonant appears in, but for the

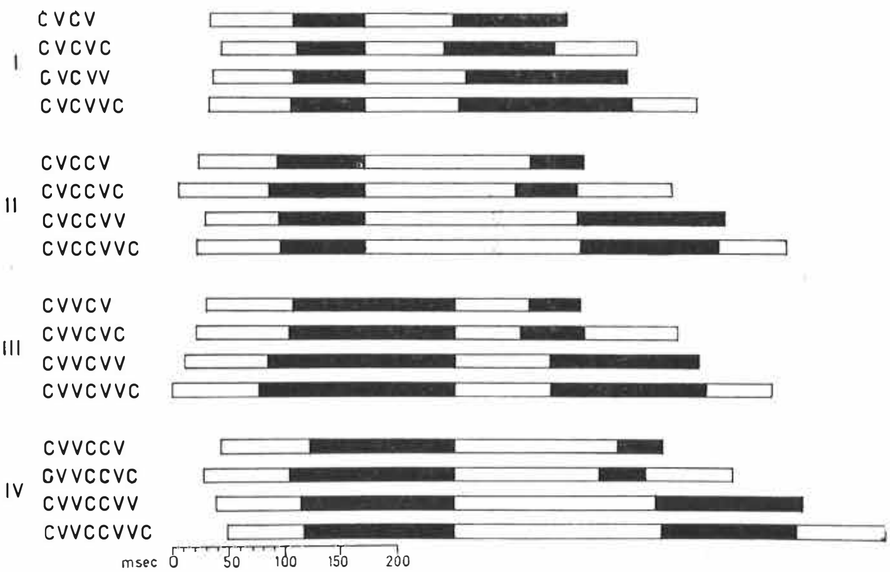


Figure 7. Graphic presentation of the durational distribution of consonant and vowel segments in different bisyllabic word structures. In the graphs consonant segments are symbolized with blank and vowel segments with black bars. The respective information is found in numeral form in table 14.

same reasons (which effect the means of initial consonant duration) one is also unable to draw conclusions concerning the various structural distinctions.

In the first group in figure 7 in which a single vowel is followed by a single consonant, the duration of the first vowel (V_1) is about the same regardless of the structure of the second syllable. The durational differences between the intervocal consonant segments (C_2) are also fairly small: the intervocal single consonant is slightly smaller in duration in structures where the second vowel segment (V_2) is a single vowel. The duration of the V_1-segment seems to be constant in group II, in which a single vowel is followed by a double consonant. The duration of the vowel segment is on average slightly longer than in group I, but the structure of the second syllable in group II does not seem to influence the phonetic duration of the single vowel of the first syllable. In the durations of the intervocal consonant segments the influence of the second syllable vowel in group II is slightly clearer than in group I: the consonant segment is clearly longer in duration preceding the long V_2-segment than before the short V_2-segment. The same feature can be seen in the duration of the intervocal consonant segments in groups III and IV: a consonant segment preceding a long vowel is longer than a consonant segment preceding a short vowel. Group III is formed by structures where the V_1-segment is a double vowel and the C_2-segment a single consonant. In this group the length of the V_2 seems to affect the duration of the V_1:when the vowel of the second syllable is a double vowel, phonetically long, V_1 is longer in duration than before a short V_2. In group IV where the double vowel is followed by a double consonant, the duration of V_1 again seems independent of the length of V_2. However, the same influence of a following vowel can be observed in the C_2-segments as in the previous ones: a consonant is clearly longer in its duration preceding a long vowel than preceding a short vowel.

In the compared structures the durational distinction between a single and double vowel of the first syllable as well as between an intervocal single and double consonant is clear. A single consonant is in all structures clearly shorter in duration than an intervocal double consonant in any structure. In groups II–IV the durational distinction between segments representing single and double second syllable vowels is also clear and is even considerably greater than in the first syllable. However, in group I where there is a vowel after the /CVC/ sequence, a phonemically single vowel is in phonetic duration twice the second syllable single vowel of other structures, and only about one third shorter than a segment representing a double vowel in a corresponding position. The 'half-long' duration of a single vowel appearing after a short open syllable, which must be taken as one of the characteristic features of the form of Standard Finnish under consideration and of several Finnish dialects, will be dealt with in detail further on.

TABLE 14. Mean durations ($\overline{X}$) and standard deviations (s) of sound segments in bisyllabic word structures

word structure CVCV				word structure CVCCV			
sound segment	$\overline{X}$ (msec)	s	N	sound segment	$\overline{X}$ (msec)	s	N
C	72	22	249	C	70	22	264
V	65	14	296	V	80	16	266
C	78	23	296	CC	146	36	266
V	102	18	296	V	47	11	266

word structure CVCVC

sound segment	$\overline{X}$ (msec)	s	N
C	67	22	267
V	62	14	293
C	70	24	293
V	100	21	293
C	74	19	254

word structure CVCCVC

sound segment	$\overline{X}$ (msec)	s	N
C	80	21	165
V	87	18	179
CC	134	39	179
V	55	13	179
C	88	20	149

word structure CVCVV

sound segment	$\overline{X}$ (msec)	s	N
C	71	24	238
V	65	16	282
C	89	26	282
VV	144	28	282

word structure CVCCVV

sound segment	$\overline{X}$ (msec)	s	N
C	66	22	279
V	78	18	279
CC	188	36	279
VV	131	18	279

word structure CVCVVC

sound segment	$\overline{X}$ (msec)	s	N
C	73	23	218
V	67	16	237
C	83	26	237
VV	154	26	237
C	59	16	217

word structure CVCCVVC

sound segment	$\overline{X}$ (msec)	s	N
C	73	31	133
V	77	18	174
CC	192	40	174
VV	123	22	174
C	60	19	152

words structure CVVCV

sound segment	$\overline{X}$ (msec)	s	N
C	75	23	129
VV	145	29	142
C	67	19	142
V	46	10	142

word structure CVVCCV

sound segmend	$\overline{X}$ (msec)	s	N
C	80	28	70
VV	130	26	80
CC	145	27	80
V	42	9	80

word structure CVVCVC

sound segment	$\overline{X}$ (msec)	s	N
C	82	23	81
VV	148	32	99
C	58	20	99
V	57	14	99
C	85	18	98

word structure CVVCCVC

sound segment	$\overline{X}$ (msec)	s	N
C	76	20	17
VV	149	23	18
CC	129	22	18
V	41	9	18
C	78	18	13

word structure CVVCVV

sound segment	$\overline{X}$ (msec)	s	N
C	72	27	118
VV	168	28	119
C	85	28	119
VV	133	23	119

word structure CVVCCVV

sound segment	$\overline{X}$ (msec)	s	N
C	75	20	108
VV	138	28	108
CC	181	29	108
VV	132	16	108

word structure CVVCVVC

sound segmend	$\overline{X}$ (smec)	s	N
C	77	26	85
VV	175	29	90
C	87	30	90
VV	140	21	90
C	59	17	90

word structure CVVCCVVC

sound segment	$\overline{X}$ (msec)	s	N
C	68	22	54
VV	136	28	60
CC	185	29	60
VV	123	19	60
C	59	20	50

3.7.6. General review of three-syllable structures

In figure 8 on page 130 and in the table 15 the average durations of sound segments in the three syllable structures included in the study material are presented. As in the general review illustrating the durational distributions of bisyllabic structures, the words here, whose durations are used in the calculations of the means, do not form minimal pairs between all structures either. A different sound composition in words representing various structures can affect the means to a certain extent, so there is no reason to draw comparisons based on these means in order to find the durational differences of individual sound segments between structures. It can be proved, however, that the durations of segments representing single and double sounds correspond approximately to segment durations of the bisyllabic structures. The durations of single V_1 and single C_2 also show least variance in the three syllable structures. The longer duration of V_1 in the CVCCVCCV-structure is caused above all by the fact that there are only fricatives and nasals, which have a lengthening influence on the preceding vowel, as double consonants following vowels in the sample words of the structure.

A half-long second syllable vowel appears in three syllable structures under the same conditions as in the two syllable ones, after the /VC/-sequence. In the CVCVCV-structure V_1/V_2 is on average about 1 : 1.3, but the corresponding ratio in the CVCVCVV-structure is 1 : 1.6. According to the total means the single V_2 of the CVCVCVV structure and the double V_2 of the CVCVVCV-structure would be almost of the same duration, but when the structure is made with only minimal word pairs the double V_2 of the CVCVVCV-structure turns out to be almost 1.5 times the duration of the single V_2 of the CVCVCVV-structure. In the same way as the single vowel of the second syllable the single vowel of the third syllable is also longer in duration after a short syllable than after a long syllable. The durational differences between single vowels of the third syllable are, however, smaller than between single vowels of the second syllable.

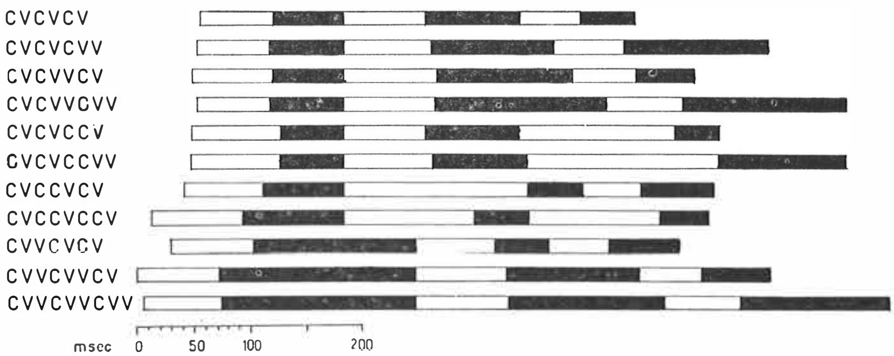


Figure 8. Graphic presentation of the durational distribution of the consonant and vowel segments in a set of three-syllabic word structures. The respective numeral values are presented in table 15.

TABLE 15. Mean durations ($\overline{X}$) and standard deviations (s) of sound segments in a series of 3-syllable word structures

word structure CVCVCV

sound segment	$\overline{X}$ (msec)	s	N
C	64	19	193
V	64	13	216
C	71	21	216
V	84	15	216
C	54	16	216
V	49	12	216

word structure CVCVCVV

sound segment	$\overline{X}$ (msec)	s	N
C	64	20	30
V	69	13	30
C	77	25	30
V	110	16	30
C	64	12	30
VV	129	14	30

word structure CVCVVCV

sound segment	$\overline{X}$ (msec)	s	N
C	71	21	51
V	65	13	60
C	81	31	60
VV	121	22	60
C	57	13	60
V	52	11	60

word structure CVCVVCVV

sound segment	$\overline{X}$ (msec)	s	N
C	65	16	39
V	68	13	49
C	80	29	49
VV	153	27	49
C	68	11	49
VV	147	20	49

word structure CVCVCCV

sound segment	$\overline{X}$ (msec)	s	N
C	79	25	50
V	56	13	50
C	71	23	50
V	84	16	50
CC	139	33	50
V	40	9	50

word structure CVCVCCVV

sound segment	$\overline{X}$ (msec)	s	N
C	78	17	50
V	58	11	50
C	78	16	50
V	84	16	50
CC	171	24	50
VV	115	10	50

word structure CVCCVCV

sound segment	$\bar{X}$ (msec)	s	N
C	69	21	149
V	75	15	160
CC	161	45	160
V	49	11	160
C	53	16	160
V	65	15	160

word structure CVCCVCCV

sound segment	$\bar{X}$ (msec)	s	N
C	81	16	47
V	91	16	47
CC	116	29	47
V	50	14	47
CC	116	33	47
V	44	9	47

word structure CVVCVCV

sound segment	$\bar{X}$ (msec)	s	N
C	72	25	69
VV	143	24	99
C	69	25	99
V	48	10	99
C	53	16	99
V	63	16	99

word structure CVVCVVCV

sound segment	$\bar{X}$ (msec)	s	N
C	72	22	30
VV	176	17	30
C	81	28	30
VV	120	25	30
C	55	11	30
V	61	8	30

word structure CVVCVVCVV

sound segment	$\bar{X}$ (msec)	s	N
C	69	17	28
VV	174	20	29
C	82	24	29
VV	140	18	29
C	68	9	29
VV	132	15	29

3.8. *The half-long vowel*

The phonetic durational distinction sustaining the single phoneme/double phoneme distinction is in many cases clearly segmental in character, but the timing ratios of the sound segments of certain word structures, which are minimally contrastive on the phonemic level, differ to such an extent that phonetic distinction cannot be adequately described just with the help of the durational difference in a certain sound segment. Description of phonemic quantity by means of the parametre of segmental duration is mostly complicated by the appearance of the so-called 'half-long' vowel as a representative of the phonemically single second syllable vowel in the (C)VCV(C) structure.[1] Definition of V_2 of a (C)VCV(C) word as half-long only means that in this structure the phonemically single V_2 is on average over one third longer than the corresponding vowel in V_1 position and as much as twice as long as phonetically single V_2 in other bisyllabic word structures (cf. Wiik 1965, p. 130). The phonetically longer duration of the half-long vowel is particularly clear in bisyllabic structures: when a single vowel is preceded by a sequence of single vowel and single consonant, the vowel is half-long, i.e. $V_1 < V_2$. If, on the other hand, a second syllable single vowel is preceded by one of the sequences (C)VVC-, (C)VCC-or (C)VVCC-, the vowel of the second syllable is phonetically short and shorter in duration than the vowel of the first syllable. The three syllable structures used in the present research seem show that the 'half-long' vowel appears in three-syllable and multisyllable words under the same conditions as in bisyllabic words, even though the absolute durational differences between vowels are smaller in the former: a phonemically single vowel is longer in duration after a VC- sequence than in the corresponding position after a VVC-, VCC- or VVCC-sequence. This holds true with regard to the duration of the second syllable vowel, as the comparison of the average durations of the V_2 segments in CVCVCV/CVCCVCV/CVVCVCV structures in figure 8 indicates. The vo-

[1] By using the term 'half-long vowel', the present author does not mean, as in the descriptions of some Finnish dialects, the subjective impression of longer duration in the V_2-segment. When, for example in the phonetic description of the dialects of South-west Finland, a single vowel following a short open syllable is marked half-long (in Finno-Ugrian transcription, for example,[takà]), it may be taken to mean the auditory impression that a vowel, particularly in a stressed position, following a short open syllable sounds stronger or more prominent than in standard Finnish or in many of the other major dialects of Finnish. The impression of longer quantity or of prominence called half-long quantity may be partly the result of longer V_2-duration, but obviously also of the manifestation of a different stress pattern in the South-west dialects or by their word–tone which differs from that of other Finnish dialects in that the pitch-peak is shifted further away from the beginning of a CVCV word. The fact that the phonetically 'half–long' vowel has been shown to be widespread even in dialects where aural-phonetic descriptions have not referred to a vowel after a short open syllable as half-long, indicates that in the perception of the 'half-long' vowel in the South-western dialects some other factor besides the absolutely longer duration of the V_2 is decisive.

wel of the third syllable is also longer in duration if it is immediately preceded by a sequence of a phonetically short vowel and single consonant. V_3 is longer in duration in the CVVCVCV and CVCCVCV structures, where the third syllable vowel is preceded (because of the phonemically double V_1 and C_2) by a sequence of a very short vowel and single consonant, than in the CVCVCCV and CVCCVCCV structures which do not fulfil this condition. The structure CVCVCV shows that the half-long duration of the phonemically single second syllable vowel also shortens the duration of the third syllable vowel as do the double consonant and double vowel preceding the third vowel. In the four-syllable CVCVCVCV structure, the alternation between short and half-long vowel duration is again just as clear as in the CVCV words. From the point of view of the syllable structures the lengthening tendency of a single vowel indicated here can be defined thus: a single vowel following a short (open) syllable tends to be longer in duration than a single vowel following a long syllable (closed syllable with single vowel and open or closed syllable with double vowel).

3.8.1. The effect of the half-long vowel on the phonetic distinctions of structures

The effects of the half-long vowel on the phonetic manifestation of the single vowel/double vowel, as well as the single consonant/double consonant distinction, in various word structures are of two kinds. On the one hand, its existence facilitates an increase in the phonetic distinction between some structure pairs, and, on the other hand, it particularly reduces the possibilities for showing the single/double vowel distinction after a short open syllable.

If the bisyllabic basic structures (C)VCV/(C)VVCV/(C)VCCV are examined, the half-long vowel makes phonetic distinction between them very effective: the presence of the half-long vowel indicates that the preceding V- and C- segments are single vowel and consonant. If, however, the V_2 is not half-long, the possibility of (C)VCV phonemisation is automatically eliminated and the choice is between three structures and is determined by the durational contrast between the vowel and consonant segments preceding the short V_2: if $V_1 > C_2$ the phonemization will be (C)VVCV, if $V_1 < C_2$, (C)VCCV, if $V_1 = C_2$, CVVCCV. By these durational contrasts between vowels and consonants is not meant the absolute duration of segments, but the relative duration whose phonetic duration value is determined by the intrinsic duration of each vowel and consonant.

Comparison of single vowel/double vowel durational distinctions in different word positions in the same phonetic frame showed that a sound segment representing a double vowel is generally at least two times as long in duration as the corresponding single vowel. However, a very clear exception in this respect is presented by structure pairs where a single vowel or double vowel appears after a short syllable. In these the duration of the double vowel is only about 1.5 times as long as the duration of the phonetically half-long single vowel. The durational ratio of the second syllable single to the second syllable double vowels in the three-syllable CVCVCV and CVCVVCV words is approximately the same (1 : 1.5).

The difference between the structures is not, however, merely contained in the different durations of the segments representating a phonemically single and double vowel. A consonant preceding a vowel and sometimes even the preceding vowel are also of significantly different duration in single and double vowel structures. Some kind of measure of the durational distinction between structures is provided by the sum of the absolute values of the durational differences between all the changing V_1, C_2 and V_2 segments. In the CVCCV/CVCCVV pair it is 125 msec, in the CVVCV/CVVCVV pair 125 msec, in the CVVCCV/CVVCCVV pair 139 msec, but in the CVCV/CVCVV pair only 52 msec.

According to Wiik's observations, (1965, p. 130–134), a half-long second syllable vowel is about 50 msec longer in duration than a corresponding short vowel (e.g. in the pair *muta* 'mud'/*musta* 'black') and a double vowel following a short first syllable is correspondingly about 50 msec longer than a double vowel following a long first syllable. In the present study, a double vowel following a short first syllable is also significantly longer in duration than one following a long first syllable, but the durational difference is merely 12 msec. There would seem to be no justification for calling the double vowel of the second syllable in the CVCVV(C) structure over-long in phonetical duration on the basis of such a slight durational difference. It would, of course, be as expected if, to maintain the CVCV/CVCVV opposition, the double vowel following the short open syllable were pronounced so long that there was adequate distinction. A vowel is indeed lengthened in speech when there is a particular reason for emphasizing that it is a long vowel that is being uttered (for example, when reading word forms detached from the context of a sentence) but in continuous speech the durational difference tends to be levelled out.

The reason for the appearance of the half-long vowel and for the levelling out of the related second syllable V/VV distinction may be sought in various ways. In the word structures of the present-day language, the half-long vowel is not appropriate because of the weakening of the second syllable V/VV distinction. It is also a feature which breaks the system of segmental manifestations of quantities. In the distinction between different structures, a second syllable single vowel equally short in all cases, even after a short syllable, would accord most with expectations. This, however, is not the case. In addition to the standard form of the language illustrated here, Finnish dialects also widely manifest the second syllable vowel as clearly longer in duration than the first syllable vowel. But even in dialects with fundamentally different durational relationships (the Isokyrö-Hamina group in the table in Wiik &Lehiste 1968, p. 571) the V_2 of a (C)VCV word is generally longer in duration than, or approximately as long as, the first syllable vowel.[1]

3.8.2. Diachronic survey

In so-called proto-Finnic there were both short and long vowels in the first syllable of a word, but in the following non-stressed syllables there were only phonemically short vowels and perhaps diphthongs of short value. There may have been both single consonants and consonant combinations on the syllable boundary – also geminates. So the structures (C)VCV(C), (C)VVCV(C), (C)VCCV(C)etc. are presumed to appear in proto-Finnic, but not, for example, the structure (C)VCVV (Hakulinen 1961, p. 34; Rapola 1966, p. 24). Since the phonetic duration of the second syllable vowel did not have a distinctive function, it could have varied freely and might have been 'half-long' in the (C)VCV structure as in present-day Estonian and – partly – Finnish. If the second syllable was really half-long, the phonetic contrast between quantity structures of the lan-

[1] According to Wiik and Lehiste, V_2 is considered half-long if the V_1/V_2 ratio is more than the arbitrarily selected ratio 1:1.3. Wiik has continued to map out the dialect georgaphical appearance of the half-long vowel measured in the natural speech of selected dialect guides. The results have confirmed and made more accurate earlier results (Wiik and Lehiste 1968) based on the speech of university students. References in the present study to the manifestation of 'half-long vowels' in Finnish dialects are partly based on preliminary results of Wiik's study.

guage form were quite effective: in the CVVCV and CVCCV types V_1 and C_1, respectively, have been phonetically long and the V_2 segment phonetically short, whereas in the CVCV structure V_1 and C_2 have been phonetically short and V_2 half-long. The half-long vowel would have been a very suitable feature in a language form in which the duration of the C_2 segment could not be used as support for the quantity distinction of the first syllable vowel as, for example, in Swedish.

However, the position of the half-long vowel in the quantity system was radically altered when long vowels developed in the following syllables as a results of the disappearance of consonants from between single vowels (e.g. *kalaTa>kalaa). Important morpheme functions, which were earlier supported by whole syllables, were now only represented by the lengthened duration of a vowel. Next to the functionally single but phonetically lengthened vowel of the second syllable there came another one which was also longer than normal short, but functionally double. In system changes such as the disappearance of single consonants in later syllables it is to be expected that there will be gaps in the arrangement, structure pairs whose phonetical distinction does not fulfil the new requirements. Such has been the distinction between single vowel and double vowel following a short open syllable. In dialects in which the CVCV and CVCVV structures were not allowed to assimilate (e.g. [kala·] = {kala} or {kalaa}) there have been various ways in which a sufficient difference between structures has been preserved. For some reason, however, the duration of the new, phonemically long vowel was not lengthened so much that the difference between structures thus became perceptually adequate. It would be tempting to imagine that the reason has been the restriction set by perception: insufficient ability to distinguish between two kinds of long duration merely on the information of one segment. For example, in Estonian, the only language known to have three phonemic 'degrees of quantities', the functional quantity distinction of a stressed syllable requires for its domain a sequence of two syllables in which only the ratio between the second syllable vowel and the vowel or consonant of the first syllable determines the phonological interpretation of the duration of the sound preceding it.

Instead of lengthening the V_2 segment, it has been possible in Finnish dialects to lengthen a little the V_1 segment of the CVCV structure and shorten the V_1 segment of the CVCVV structure and in this way to increase the contrast between these structures. (According to the results of Laurosela's

(1922) measurements this should happen within the area of the dialect of South Ostro-Bothnia – the same solution is also perhaps applicable in certain dialects of Häme-Satakunta). The third possibility has been to use the tendency, which was perhaps already in existence, of a consonant to lengthen in duration before a long vowel. The conscious increase in this lengthening led, however, to the weakening of the distinction between a short and long consonant preceding a new long vowel: the [CVC·V:] structure appeared along with the structure [CVC:V:], and finally the C_2 of the CVCVV structure became equal to the original geminate, causing so-called general gemination of the consonants.

3.9. *The effect of the word length on the duration of sounds*

There are various references in linguistic literature to the phenomenon called isochrony or equalization, which appears in the way in which, for example, the duration of an individual sound depends, apart from other factors, on the length of the word in which it appears. It has been pointed out that the duration of the vowel /a/ gets shorter as the word lengthens in the French series *pate-pâte-patisserie*; or the duration of the vowel /i/ in the Swedish word series *fisk-fiskare-fiskarstuga* (Malmberg 1944, p. 10–11; 1967, p. 150; Abercrombie 1967, p. 97). From this kind of observations Collinder (1964) was led to assume, that there exists a certain tendency, of restricted manifestation in each language keeping the standard word duration the same. According to Collinder the unit which regulates the duration of an individual sound in such series as the ones mentioned above, is the word, »a sequence of sounds which represents a unified concept».

Eli Fischer-Jørgensen (1964), in her analysis of Danish durations, obtained results which seem to contradict the theory that duration of a single sound in a word is in inverse proportion to the number of sounds (or syllables) in the given word. Maack (1949, p. 200) has also shown that the durational ratio of the stressed vowels in mono- and bisyllabic words in German is accidental: »The stressed vowels in one syllable words are often longer, but just as often somewhat shorter than in bisyllabic paroxytones – in extensive material they are more or less of exactly the same length.» Therefore according to Fischer-Jørgensen it is doubtful whether the word is really relevant as the unit which effects duration: »In the cases where measurements have been made on isolated words, there is a coincidence of word, stress group and intonation group, and it may just as well be one of the higher units that is relevant. There may also be differences in this respect according to the structure of the particular language» (Fischer-Jørgensen 1964, p. 200–201 cf. also Harms 1968, p. 408). Elert (1965, p. 138) has criticized the appearence of the isochrony principle at the word level even when sentence frames are used in research instead of separate words. The identical sentence frames or structurally similar frame sentences used by phoneticians may conceal the normal temporal patterns which exist in fluent speech. According to Elert, in normal free speech the intervals between the stressed syllables in utterances where polysyllabic words occur are really longer in normal free speech than in a series of shorter words. Even in Swedish word length or the number of syllables effects the durations of individual sounds but the durational differences, at least between first syllable vowels in one and two syllable words, are fairly small. The duration of the stressed long vowel in Swedish in the Accent I word in bisyllabics spoken in a sentence frame is 93 % and in the Acc II words 83 % of duration of monosyllabics. In separately read words the durational differences are to some extent greater: the duration of the stressed long vowel of the two syllable Acc I word is 85% and of Acc II 78 % of the duration of the vowel of a one syllable word. The durational ratios of short vowels are about the same as those of long ones. The results of Lindblom's study (1968), based on nonsense words read by Swedish people, differ from those obtained by Elert. According to Lindblom (1968 fig 1–A–1) a stressed long vowel in a bisyllabic word is only about 60 % of the duration of a long vowel in a one syllable word. The stressed long vowels of three and four syllable words are about the same in duration and shorter than the stressed vowels of a bisyllabic word. The vowel of a five syllable word is again slightly shorter than the vowel of three and four syllable words. The effect of the number of syllables in a word, according to Lindblom's table is the same in short vowels, even though the absolute differences in duration are slightly smaller.

The effect of word length or the number of syllables on the duration of individual sounds has also been illustrated in several experimental tests on sound duration in Finnish and in other languages related to Finnish. For example, Meyer & Gombocz (1907–08, p. 141) have shown that in Hungarian »die Dauer der Vokale mit zunehmender Anzahl der zu einem Sprechtakt vereinigten Silben stetig abnimmt». Shortening is most obvious when one compares the short vowel of a one syllable word to the short first vowel of a bisyllabic word (in Meyer and Gombocz's examples 44 %), while in the corresponding long vowels the shortening is only 11 %. In Meyer and Gombocz's examples the shortening in comparisons between one syllable structures and two, three, four and five syllable structures is about the same in size and varies in different comparison pairs from 0 to 15 % in both the consonant and the vowel segments.

The same phenomenon, shortening of the duration of sounds as the number of syllables increases, has also been illustrated in Lappish by Äimä (1914, p. 204). »A certain sound quantity is in general somewhat longer in duration in a bisyllabic word than the same quantity in the same position in a 3-syllable word, and in this latter longer still than in a 4-syllable word.» Äimä uses the term *Rhytmischer Abstufung* to classify this phenomenon.

The lengthening of both vowels and consonants in relation to an increase in the number of syllables of a word has also been observed by Donner in his study of quantity relations in the Salmi dialect (1912, p. 9 and general review p. 33). Laurosela demonstrates, in his study on the dialect of Ostro-Bothnia that the vowels of open first syllables in 1-, 2-, 3- and 4-syllable words are related to each other thus: 1 : 0.78 : 0.57 : 0.53. Alongside these relations Laurosela presents the values differing from those above of the relationships between short initial syllable vowels: the ratio of short vowels of the 1st syllable to two and three syllable words is 11.1 : 9.1, the corresponding ratio of three and four syllable ones 9.1 : 8.3 (Laurosela 1922, p. 164).

Sovijärvi has also illustrated, in his study on the Ingrian dialect of Soikkola (1944, p. 11), a phenomenon affecting the duration of these sounds. According to Sovijärvi »together with an increase in the number of syllables of a phoneme or a lengthening of syllables the durations of sounds correspondingly decrease.» By »phoneme» Sovijärvi means here a word, or an expression longer than a word, which is, for example, spoken all at one blow onto the kymograph.

The shortening of the durations of individual sounds is quite clear in Standard Finnish according to Palomaa's comparisons (1946, p. 49) read onto the kymograph by one person. But it is precisely Palomaa's samples, in which the speech tempo seems to have been very slow throughout, which give reasons for wondering whether the factor affecting durations is perhaps not word length but the length of sequence forming the expression read onto the kymograph the: shorter the sound sequences read in one breath group the slower the tempo at which the readers pronounce them. It is worth noticing that in neither Meyer and Gombocz's nor Sovijärvi's observations was the word given as the unit influencing sound duration, but 'Sprechtakt' and 'phoneme' i.e. the utterance. Equalisation or quantitative congruence, as Sadeniemi (1949, p. 105) calls the phenomenon, would appear, therefore, within the rhythm units and not within the phonemic words. There would also be a prevailing tendency in Finnish to space out the intensity peaks or stressed syllables of speech at regular distances from each other (Sovijärvi 1957, p. 320).

When the frame sentences of the analysis material in this study are so arranged that the words under examination form a speech measure (see chapter 3.12) of their own in each case, one could expect the durational differences to become manifest in comparisons between the durations of the sounds in corresponding positions in longer and shorter words. In table 16 there is a pair comparison of sound durations in CVVCV-and CVVCVCV, CVCCV- and CVCCVCV-type, CVCV- and CVCVCV- and CVCVCV and CVCVCVCV type structures which are identical in their common parts. The statistical significance observed in differences between durations is also presented. The differences are also presented graphically in figure 9.

TABLE 16. The effect of the word length on sound duration. In the first three sets the magnitude and significance of the difference of the average durations of the corresponding phonemes of the two or three syllabic words identical in their common parts are compared with the help of the t-test, (e.g. *laaka*'flat'/ *laakana*'as flat'). In the fourth set the durations of the sound segments of three syllable words and correspondingly, four syllable words, are compared (for example *takana* 'behind' / *takanani* 'behind me').

Compared structures
CVCCV – CVCCVCV

Soune segment	Structure CVCCV Mean (msec)	n	Structure CVCCVCV Mean (msec)	n	Difference ofmeans
C_1	69	138	68	138	$+1^{non}$
V_1	76	139	76	139	0
C_2	151	139	158	139	$-7^{non}(p<20\%)$
V_2	47	139	50	139	$-3^{non}(p<10\%)$
C_3			52	139	
V_3			65	139	

Compared structures

CVCV – CVCVCV

Sound segment	Structure CVCV Mean (msec)	n	Structure CVCVCV Mean (msec)	n	Difference of means
C_1	73	160	65	173	+8****
V_1	65	198	64	196	$+1^{non}$
C_2	79	198	72	196	+7****
V_2	102	198	84	196	+16****
C_3			53	196	
V_3			50	196	

Compared structures

CVVCV – CVVCVCV

Sound segment	Structure CVVCV Mean (msec)	n	Structure CVVCVCV Mean (msec)	n	Difference of means
C_1	71	70	72	59	-1^{non}
V_1	146	80	145	79	$+1^{non}$
C_2	71	80	71	79	0
V_2	48	80	49	79	-1^{non}
C_3			53	79	
V_3			65	79	

Compared structures

CVCVCV – CVCVCVCV

Sound segment	Structure CVCVCV Mean (msec)	n	Structure CVCVCVCV Mean (msec)	n	Difference of means
C_1	56	30	64	30	-8^{non}(p 10%)
V_1	63	30	66	30	-3^{non}
C_2	86	30	94	30	−8*
V_2	80	30	91	30	−11****
C_3	48	30	53	30	-5^{non}
V_3	55	30	60	30	-5^{non}
C_4			52	30	
V_4			73	30	

The comparisons do not altogether support the theory that a given sound is significantly shorter in a longer word than a corresponding sound in a shorter word. For example, the first syllable vowel is almost exactly the same duration in the word types *taka*, *takana*, *takanani*. In the structure pairs *takka / takkana* and *laaka/laakana* there is no significant difference and even the durational differences which can be interpreted as indicative support the three syllable structure. On the other hand in the durations of the sound segments of words formed with CV-syllables there are some very strongly significant differences which, however, are not purely related to the length of the word: in the CVCV-structure, C_2 and V_2 are significantly longer than in the CVCVCV-structure, but in the four syllable CVCVCVCV-structure C_2 is almost significantly and V_2 strongly significantly longer than in the shorter CVCVCV-structure. Two and four syllable structures cannot be compared directly here because the word groups used in comparisons between two and three syllable structures and in comparisons between three and four syllable structures differ from each other. Neither is there any reason for paying attention to differences between the C_1 segments, because they are also affected by the different final sounds in the previous words of the frame sentence.

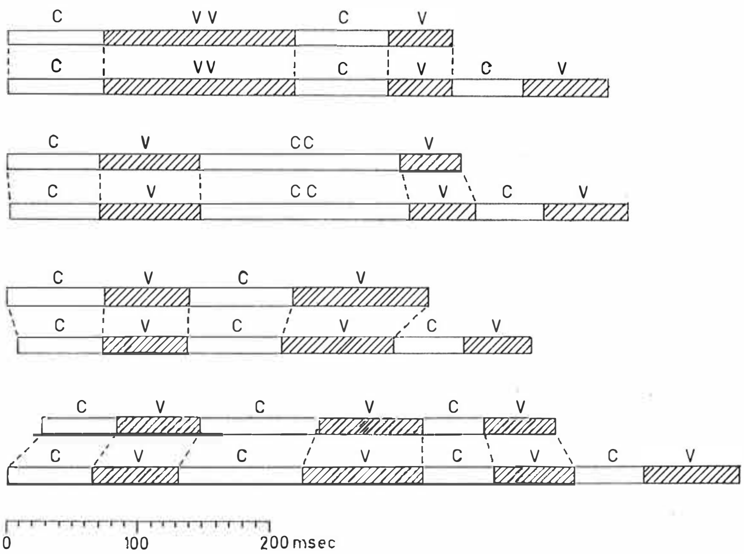


Figure 9. The effect of word length on the duration of sounds. Comparison of the durations of the sound segments of three different bisyllabic and one three syllable word structures with structures having identical common parts to which one CV-syllable is added in each case (for example *sata* 'hundred'/*satama* 'port': *takana* 'behind'/*takanani* 'behind me').

It is clear that durational differences appearing in the structural pairs CVCV/ CVCVCV and CVCVCV/CVCVCVCV are not related to the different number of segmental sounds in various structures in such a way that the absolute duration of a certain sound segment would be shortened together with an increase in the number of sounds, because the ratio in the pairs *takana/takanani* is precisely the reverse. Even though, as far as sentence rhythm is concerned, the CVCVCVCV-word only forms one weak-stressed rhythm unit, or speech measure, in its frame sentence, in articulation timing it is clearly divided into two identical CVCV-measures in each of which the first CV-syllable is heard as word-stressed in Finnish (cf. Lehiste 1965 B, p. 180, footnote).

The durational relationships of three, four, and five syllable CV-words given by Lehiste (1965 B, p. 180) which are based on the speech of one native Finnish speaker correspond very well to the results of this study. The symmetry of the durations in the CVCV-sequences is very clear in the speech of Lehiste's informant, so much so that the means of the durations of the V_1 and V_3, as well as of the V_2 and V_4 segments, in the eleven occurences of the CVCVCVCV word *manalana* were precisely the same (8.2. and 11.5 sec).

One may, therefore, state that word lenght in Finnish has no significant effect on the duration of the individual sound when the words appear in the same sentence stress position. At least one cannot consider that the duration of the sounds of the initial syllables of a word would function as an indication of the length of the word so that shorter duration of the sounds of the initial syllables would act as a cue of polysyllabicity of words. Comparison between word lengths in the present study has not, however, extented to monosyllabic words which play a fairly small part in Finnish vocabulary and which are difficult to put into a sentence stress position corresponding to longer word structures (c.f. chapter 3.2.). According to Wiik (1965, p. 119), the stressed vowels are about 18 % longer in monosyllabic than in bisyllabic words. In the identical sentence frames »Mitä sana --- tarkoittaa» 'What does the words --- mean?' and »Mitäs --- merkitsee?» 'What does --- mean?', used by Wiik, the difference noted in the words read is, however, so small compared to, for example, the corresponding durational differences in English, that there is no reason to consider it a characteristic feature of the Finnish durational system[1]. Obviously the ratio also varies according to

[1] It is unlikely that the duration of the initial vowel in one and two syllable words, for example, would be different when the words appear in a similar stress position and in a similar speech

the kind of two syllable structure used in comparison with the corresponding vowel duration of a one syllable word.

3. 10. *The effect of stress on sound duration*

It is a known fact in many languages that stress affects the duration of a vowel in that a vowel is clearly longer in a stressed than in an unstressed position. The duration of a vowel can then be a vital cue, in perception, of the stress of a syllable. On the basis of the original material in this study it is not possible to state directly the modifying effect of stress on the sound durations in Finnish nor can one show the significance durations or durational relations have in the perception of stress.[1] The test words of this study were in such positions in their frame sentences that they formed a speech measure of their own in the sentence rhythm, but did not receive the main sentence stress. In content the sentences were so constructed that the word to be examined did not receive prominence, or emphatic stress, differing from normal sentence stress. This solution was reached to distinguish the timing features indicating the phonemic composition of words and the phonetical parameters of stress from each other.

In order to show the part played by duration in sentencestress perception one should first be familiar with the normal durational patterns in relatively unstressed words. If stressed words were to be examined, it would still be unclear which of the features of duration, intonation and intensity belong to the parameters of the sentence-stress. Nor is the character of sentence-stress self-evident: some word perceived as stressed may contain phonetical cues causing prominence, but perception of 'stress' in a word may also be caused primarily by its syntactical position or by its logical relation to the total content of the utterance.

Preliminary indications of the part played by durational changes among the parameters of stress were given by a separate test in which one informant read a number of sentence pairs. In the first of these the test word is clearly in the sentence stressed position, in the latter it is in a relatively unstressed position. The sentences were *Onko sana* ___ *(primary-stressed) kirjoitettu oikein* 'Is

measure (for example, in the pair *tuo ja vie* 'he brings and takes'/*tuoja vie* 'the bringer takes' or (*Kai saa katsoa* 'Kai may watch'/*Kaisaa katsotaan* 'they are watching to Kaisa').

[1] Sovijärvi (1958) has analyzed the perceptual parametres of Finnish word-stress with a kind of tape cutting method but he has overlooked the possible effect which the shortened duration of the stimulus in itself may have on stress perception.

the word ___ written correctly' and *Minun mielestäni* ___ *(non-primary-stressed) kirjoitetaan juuri näin* 'In my opinion ___ is written just like that'.

In (C)VCV and (C)VCVV type test words all the sounds, in a stressed position, with the exception of C_1, were longer than in an unstressed position. The lengthening was on average 8 % in the V_1 segment, 20 % in the C_2 segment and 32 % in the V_2 segment. The relative lengthening was almost exactly the same in both the (C)VCV and the (C)VCVV structures. The number of stressed/unstressed pairs measured was 14. In the structures CVVCCV, CVVCCVV, CVVCV, CVVCVV, CVCCV, CVCCVV (number of pairs 17) most lengthening in sentence stressed position occurred in double vowels (33 %), the intervocal single and geminate consonants lengthened in about the same proportion (about 25 %) and the least lengthening occurred in single vowels (15 %). The lengthening of 1st syllable double vowels, according to the limited material, is clearly less than the lengthening of 2nd syllable double vowels (15 % in the 1st syllable and 50 % in the 2nd syllable).

This preliminary comparison, with the use of restricted material, of sound duration in words with strong sentence stress and weak sentence secondary stress is clearly significant: apart from other phonetic factors (above all pitch) a clear increase in the duration of sounds in a stressed word (or measure – but not a mere syllable) and a certain change in durational pattern seems to be related to sentence stress. This preliminary comparison also seems to indicate that the segmental manifestation of stress does not affect only, or even mostly, the durations of the sound segments of the syllables which are said to have word-stress in Finnish, but the durations of the whole measure formed by two syllables In the compared word structures the increase in the duration of vowels in syllables said to have word-stress was, on the contrary, clearly less than in 2nd syllables which are word-unstressed. The reason for the lengthening of segments following a first syllable vowels can be seen to be a result of the vowel duration of the first syllable of a word being more bound to phonemic function than the duration of the following segment: the lengthening of the duration of the first syllable single vowel would change the contrast of vowel and consonant segments in such a way that it would be identified as a double vowel. On the other hand the following segments of the (C)VCV(C) sequence forming a measure are in a better position to accept an increase in quantity caused by stress. Accordingly one would expect that, under the influence of sentence stress, a first syllable double vowel, whose increase in duration does not affect the meaning of a word, would lengthen more than following segments, but at least when it is followed by a long consonant or a long second syllable vowel the increase in the duration of following segments is greater than the durational increase of the first syllable vowel.

3. 11. *Durational distribution in slow speech*

In very slow speech the durational relations of segments of the word structures do not stay the same as in normal speech, but there is a certain regularity in the lengthening of the segments. Kierimo (1929) has made comparisons between the sound durations of words read separately into the kymograph and words in a weak-stressed position in a sentence frame. According to Kierimo shortening in words in a sentence environment is most noticeable (46 %) in intervocal geminate consonants. The shortening of the final vowel of a word is 37.6 % and the shortening of a first syllable vowel only 17.3 %. The smallest difference, according to Kierimo's observations, is in the durations of the sounds of a CVCV word read within a sentence and separately. In these the average shortening is 18.3%.

In connection with the present study the ten speakers were asked to read a number of separate words at a very slow speech tempo. The lengthening of the duration of the V_1, C_2, and V_2 segments of these words is shown in table 17 as percentages compared to the sound durations of the corresponding words read in a sentence environment.

The lengthening of the first syllable single vowel is on average clearly less than lengthening of the first syllable double vowel. The lengthening of the intervocal single consonant is also less than in the corresponding double consonants. Lengthening in the consonant cluster is clearly greater in the first member of the cluster. However, the calculated standard deviation of the measured durations of the first member of the cluster is four times the deviation of the latter member. A single vowel following a long syllable increased in duration by over 200 % and reached the same absolute duration as the single vowel of the first syllable. The second syllable double vowel also lengthened slightly more than the first syllable double vowel and the durational differences between the sounds of a stressed or unstressed syllable in speech of normal tempo have been eliminated. For example the V_2 of the CVCCV word grows longer than

TABLE 17. Percentual lengthening of sound segments in isolated wordss poken in slow tempo compared with the segmental durations of respective words spoken in the sentence frame at normal speech tempo. Every percentage is counted on the basis of the means of ten measured items.

	lengthening of sound segment				
word	V_1	C_2	V_2	C_2x	C_2y
	%	%	%	%	%
napa	105	110	104		
kasa	174	122	100		
napaa	96	81	135		
kasaa	109	100	159		
nappa	54	217	265		
kassa	80	172	198		
nappaa	66	158	105		
kansa	110	255	180	340	159
natsa	84	204	225	294	126
natsaa	83	148	158	218	68
natsan	73	185	198	264	95
vaaksa	162	188	248	230	145
vaaksaa	144	132	194	195	81
saastat	156	177	229	283	112
saastaa	140	88	168	129	64

the V_1, and in the CVCV/CVCVV distinction the V/VV ratio increases because of the considerable lengthening of the double V_2, even though the 'half-long' vowel (the durational contrast $V_1 < V_2$) of the CVCV-structure is maintained. It is possible that in slow enough speech the duration of each syllable or sound is consciously regulated by a separate command which changes the automatic patterns of natural tempo.

3.12. *Units of temporal organization of sound sequences*

The forcgoing comparisons of various word structures were made without the assumption that the word would be the relevant unit of durational distribution of sound segments. The word is, however, traditionally interpreted as the reference frame of changes in durational patterning just as in the general gemination of consonant in Finnish dialects. The reason for consonant gemination has generally been thought to be the need to balance the intensity difference between the stressed and unstressed parts of the 'word shape' (Finnish: *sanahahmo*) caused by the lengthening of the second syllable vowel which was upsetting the balance (Rapola 1966, p. 24; Sovijärvi 1944, p. 22–). On the other hand, the gemination phenomenon of consonants in several Finnish dialects seems to be clearly related to the durational patterns and specially to the appearance of the half-long vowel: consonant doubling has occurred in those very dialects in which the appearance of half-long vowel has been recently shown by Wiik. (cf. the footnote on page 136). In order to evaluate the share of the balance and the upsetting of the balance, one must decide what kind of units form the hypothetic 'word shape' and what factors have an influence on it and on the half-length of a vowel: in other words, one must find out what kind of units influence the temporal organization of Finnish.

3.12.1. Speech measure and measure

The term speech measure (Finnish *puhetahti*) is used in the descriptions of Finnish in connection with two kinds of unit: the first is the unit which describes speech rhythm, and which could be ranked with *stress groups.* This kind of speech measure is an interval which stretches from a sentence-stressed or prominent syllable to the following sentence-stressed syllable (Sovijärvi 1946, p. 122; Sadeniemi 1949, p. 55). These kinds of rhythmical speech measures have allegedly a tendency, within the limits of one larger unit called rhythm group (Finnish *rytmijakso*), to be roughly of equal length. As can be seen from Sovijärvi's analyses of rhythm (Sovijärvi 1946, p. 126–129), this kind of rhythmical speech measure may include one or two sequences which start with word stress. The latter sort of unit, which includes the stressed syllable of a word and unstressed syllables which may follow it, is called by Sadeniemi (op. cit.) measure (Finnish *tahti*). According to Sadeniemi, the word *päällekirjoitus* 'title', for example, contains one speech measure but two measures.

It is often difficult to draw a distinction between measures and speech measures. It seems, however, possible to find the following distinction in the use of the terms: Speech measures are conditioned by sentence stress, whereas measures are not; sequences of syllables which form themselves, regardless of the position of sentence stress, into one entity and whose first syllable is said to be word-stressed, are called mere measures. Thus the measure is seen as a structural unit larger than a syllable. Measures and speech measures are often identical. For example, in the sentence *Lepän vesa taittui*[1] 'the alder shoot broke off' each word contains one measure and each measure forms its own individual speech measure. However, in the sequence *hyvänä satamana tunnettu*[1] 'known as a good port' each word forms its own speech measure, but the word *satamana* 'as a port' is divided into two measures 'sata' and 'mana' of which only the former gets a weak sentence stress, but the first syllable of both is said to be word-stressed. In the sentence *Pidit lakanana peittoani*[1] 'you kept my blanket as a sheet' there are three speech measures, but both the word 'lakanana' and the word 'peittoani' contain two measures. Thus the concepts of measure and word stress are combined: the initial syllable of each measure is word-stressed (main or secondary stressed) and the measure can be defined as an interval stretching from word stress to word stress. Normally sentence stress falls on the first syllable of a word but in the case of emphatic stress a secondary stressed syllable may get the sentence stress too, e.g. [. . . ˈsataˈmana].

3.12.2. Syllable and mora

A prosodic unit smaller than a measure is a syllable. Each measure contains from one to three syllables. The number of syllables is not, however, always the best measure of unit length. If this was used, the long and short syllabled words *takana* 'behind' and *tarkkuuttaan* 'of his accuracy', for example, would be equal in length, since both are three-syllabled. Long and short syllabic structures are, however, subjectively of different lengths and they clearly behave in different ways, even with regard to stressing possibilities: it is possible, in certain situations, to stress the second syllable following a long first syllable in a bisyllabic word, e.g. *Minä sanoin kanˈta enkä kanˈto* 'I said base not stump'; it is, however, more unnatural to stress the syllable following a short first syllable, for example **kaˈsa* 'heap'. But again the last syllable of a word containing three short syllables can be stressed in certain cases, e.g. *Sanoin kasaˈna enkä kasaˈni* 'I said as a heap, not my heap'. It is quite impossible to consciously stress **kaˈsana*.

[1] The samples are out of the test sentences of the present work

If the lengths of units are measured in mora-units, so that a syllable-initial consonant immediately preceding a vowel has the mora value 0, and the other sound segments the mora value 1 (Sadeniemi 1949, p. 103), the sequences CVCV, CVV and CVC (C) are of equal value, each having 2 moras. It was observed in Chapter 3.10. that in a CVCV word it was not the first syllable, conventionally thought of as stressed, which was the only one to be lengthened. In addition, it did not receive most lengthening. On the contrary, all the sound segments were lengthened, and V_2 most of all. The domain of stress was shown to be not the first syllable but the first sequence of two moras. For this reason, a Finnish person 'hears' the first short syllable as stressed even when the peaks of intonation, intensity and duration are on the second syllable of a word.[1] In the same way as for the stress area, the two-mora sequence, and not the syllable, can be interpreted as the smallest temporal unit of speech. In the case of long syllables, this unit becomes identified with the syllable, whereas a short syllable on its own does not form such a unit.

Obviously, in a temporal analysis of Finnish, various larger and smaller units should be distinguished, irrespective of the fact that some of them can often fall together. These are the wider rhythm sequence (*rytmijakso*), the speech measure (*puhetahti*) or stress group (distinguished according to sentence stress), the measure (*tahti*, distinguished according to word stress), the sequence of two-morae and the syllable.

If it is supposed that a word stress has some standard minimum duration or amount of energy whose domain is the two-mora group, this would also help to explain the appearance of the problematic half-long vowel. The articulation of segmental sounds should, in some way or other, be synchronized with this imagined energy pulse. Comparison of the durations of segments in various structures showed that the durations of an intervocal single consonant and of the first single vowel of a word clearly vary less in different situations than the durations of other segments, perhaps in order to safeguard the V/VV distinction in the first syllable and the C/CC distinction at the syllable boundary. We can illustrate, that in the language form under consideration a certain delay constant appears in the realization of the production command of the two-

[1] Foreigners often hear the second syllable following a short syllable in Finnish as stressed (see Sadeniemi 1949, p. 60–61; Collinder 1937 p. 113). The physical stress peak may fall on the second syllable but the relevant unit for a Finnish person is not, however, the short initial syllable of a word but the measure to which the syllable belongs.

mora group due to which the intensity maximum is only reached at the beginning of the second mora. The refore the remaining intensity must be used for the production of the second mora.[1] In a (C)VCV word the peak occurs at the beginning of the V_2 segment, in (C)VVCV and (C)VCCV words at the beginning of the second member of the double vowel or at the beginning of the first member of the double consonant. The part to be left for the second mora grows bigger as the word stress increases in intensity. According to this, the C_x segment in a CVC_xC_yV word, the V_y segment in a CV_xV_yCV word and the V_2 segment in a CVCV word would increase as the degree of stress increases.[2]

[1] Intensity as used here is quite a theoretical concept and does not indicate the physical intensity of a sound. Perhaps some motor processes and their corresponding articulatory muscular actions could be seen as parametres of this 'intensity'.

[2] In the production of diphthongs, however, there can be a tendency which levels out the division of additional quantity between the members and the transition (see Karlsson 1970).

4. PERCEPTION TESTS

In order to discover what factors in the duration or durational relations of sounds affect the identification of the sound segment as single or double, a series of identification tests was prepared whose stimuli were produced by means of the 'segmentator' equipment constructed for this purpose. Neither the testing of listener's subjective judgment of quantity degrees nor the evaluation of the naturalness of test samples was the purpose of the tests. The purpose was merely to demonstrate the influence of the changed segment duration on the listener's phonemisation or on the manner in which the listeners identify as Finnish words linguistic stimuli which have been produced by changing the duration of some sound segment of the original spoken word.

Analogical tests where the stimuli are produced by modifying the original durations of prerecorded natural utterances, have been made by, for example, Lisker (1957), Abramson (1962), Bastian & Abramson (1962), Hadding-Koch & Abramson (1966) and Huggins (1968). Actually this method does not differ from the tests with synthetic stimuli where the duration of a certain segment is varied in identical environments, made by, for example Liberman & al. (1961), Delattre & Monnot (1968), Delattre & Hohenberg (1968). The stages by which the durations of sound segments were lengthened or shortened, vary in different studies from 8 to 50 msec. In the present study 15 msec was chosen as the shortening or lengthening stage for vowel segments and 25 msec as the changing stage for consonant segments. These figures are based on the preliminary tests, according to which discrimination of the durational differences of the voiceless obstruents seemed to be less exact than that of vowels and allowed a longer transitional phase which restricted the stimuli to be tested to a reasonable number (7–18 stimuli per word).

Phoneme sequences were chosen for the tests, whose vowel and consonant segments would give as great a number as possible of existing Finnish words when doubled. For the final tests the following were chosen from several structure frames, partly arbitrarily, partly because the reading of these very words was most successful in segmentation. The basic structures used here and the corresponding existing Finnish words are shown in the following list. In changing a first syllable vowel word pairs beginning with a vowel were used instead of these.

common phonetic segments	phonemic word structures[1]							
	CVCV	CVVCV	CVCCV	CVVCCV	CVCVV	CVCCVV	CVVCVV	CVVCCVV
r+i+p+i	ripi	riipi	rippi		ripii		riipii	
m+ä+t+ä	mätä			Määttä	mätää	mättää		Määttää
m+u+t+a	muta	muuta	mutta	muutta	mutaa	muttaa		muuttaa
t+a+k+a	taka		takka	taakka	takaa	takkaa		taakkaa
l+i+k+a	lika	liika	likka	liikka	likaa	likkaa	liikaa	liikkaa
k+a+s+a	kasa		kassa		kasaa	kassaa		
p+e+s+e	pese[2]		pesse[2]		pesee	pessee		
r+y+s+ä	rysä	ryysä	ryssä		rysää	ryssää	ryysää	

The tempo of speech represented by the test words read within the word lists is clearly slower than that of the words read inside meaningful sentences in a relatively unstressed position. Because of this, apart from the absolute durations of their sound segments, the durational relations of successive segments in different word structures also differ from the calculated average durational relations of sentence list words. This makes it more difficult to compare the results of perception tests and of the original normal structures, but on the other hand the difference prevents one from drawing conclusions which are too simplified about the regularities related to quantity perception.

Cutting of the test words from those read by the ten speakers in a

[1] The meanings of the listed words: *ripi* 'birch'; *riipi* '(he) stripped off'; *rippi* 'confession'; *ripii* 'it's sprinkling': *riipii* '(he) strips off'; *mätä* 'rotten'; *Määttä* (person's name); *mätää* 'of rotten'; *mättää* '(he) scoops'; *Määttää* 'of Määttä'; *muta* 'mud'; *muuta* '(something) else'; *mutta* 'but'; *muutta* 'without anything else'; *mutaa* 'of mud'; *muttaa* 'but (part.sg.); *muuttaa* 'he moves'; *taka* 'back'; *takka* 'fire-place'; *taakka* 'burden'; *takaa* '(he) guarantees'; *takkaa* 'of the fire-place'; *taakkaa* 'of the burden'; *lika* 'dirt'; *liika* 'excess'; *likka* 'girl': *liikka* 'limping': *likaa* 'soils': *likkaa* 'of the girl': *liikaa* 'too much': *liikkaa* '(he) is limping'; *kasa* 'heap'; *kassa* 'cash-desk'; *kasaa* '(he) heaps'; *kassaa* 'of the cash-desk'; *pese* 'wash!'; *pesse* '(may not) wash'; *pesee* '(he) washes'; *pessee* '(he) may wash'; *rysä* 'bow net'; *ryysä* 'crowding' (slang); *ryssä* 'Russian' (slang); *rysää* 'of the bow net'; *ryssää* 'of the Russian'; *ryysää* 'is crowding'.

[2] The imperative mood *pese* and the potential mood *pesse* morphophonemically end with a consonant (cf the notes on pages 24 and 84) but are in isolated position without any final consonant.

sentence environment would not have been successful, because the sentence intonation and the influence of the surrounding sounds on the word within the sentence make it unnatural when isolated. And secondly the very short absolute durations of some sound segments would have restricted the possibilities for changing durations. On the other hand the use of the original frame sentences as test stimuli would have meant that the listeners were performing the recognition on the basis of redundant information in the sentence frame and not on the basis of durational relations and the durations of the sound segments of a word.

4.1. *Arrangement of the listening test*

7th grade pupils from two Jyväskylä grammar schools (Jyväskylä Grammar School for boys and Jyväskylä Grammar School for girls) were listeners in the recognition test. When students, whose native district was other than Jyväskylä town or one of its neighbouring communities, were rejected there were 42 informants left, 20 girls and 22 boys. The age of the informants varied between 17 and 20 years. Both groups of girls and boys heard the sample in their own classroom at school over two Hi-Fi loudspeakers. The possible effects of outside disturbance during the test were eliminated by the listeners listening to each stimulus three successive times. For the test the listeners were given answer forms, on which there was a column for each stimulus series. At the beginning of the test the listeners were told that they were going to hear Finnish words over the loudspeakers which they would have to identify and write down on their forms. The listeners were not told the purpose of the test nor how the words heard in the test had been produced. They were, however, told that some words may sound odd, 'because they were isolated from speech', and that their task simply was to identify each word in turn. They were further told that 'in some series the same word may be repeated a few times, but will then suddenly change to something else'. The various word series were recorded in random order on a tape recorder in periods of about 15 min. After each 15 min. period there was a short pause and after every 45 mins there was a longer pause. The whole identification test which was administered in one session took three hours. Even though the test was quite strenuous and tiresome, the listeners were able to concentrate on their listening task to the very end.

The order of the stimuli within the sample series was not changed haphazardly, but the samples were heard in the original order of shortening – or lengthening – so that a certain sound segment was lengthened or shortened sample by sample from the beginning to the end of the series. Several listeners did not, however, get any idea during the whole series of what changed in each group from one sample to the other and did not, there fore, notice that the duration of a given sound was regularly changed. The reason for this may partly be that the 5 sec pause following each three successive repetitions of a stimulus and the ordinal number of the following stimulus spoken on the tape which was also followed by a couple of seconds pause, prevented the contrasting of the auditory impressions of successive stimuli. In order to find out whether the order of performance had any influence on the identification of the words, one series was given in two reverse orders. The results of these tests can be seen in figure $10._6$ on page 165. The circles joined by an unbroken line indicate the percentage distribution of the listener's responses when the second vowel segment of the original word *mätää* 'of a rotten one' is shortened, the triangles joined by a broken line show the listener's answers to the same sample series given in reverse order. The differences between the answers, which appear only at the extremities of the curve, are quite small. The centre part of the answer curves, 50 % degrees, are almost exactly joined. The differences at the extremity of the shorter vowel durations are very small, but the shortening of long duration has caused a slightly later change in perception than the corresponding lengthening of shorter durations. The shifts of identification caused by shortening and lengthening of duration are, however, so similar that the order in which the stimuli are presented does not have to be considered as having any significance in the test results. The possible effect of the order of performance of the samples has, however, also been eliminated so that a part of the sample series was given in such a way that the sounds to be changed were lengthened sample by sample, and another part was given in such a way that the sounds were shortened from the beginning of the series. On the basis of these tests it cannot be calculated whether the internal random order of the series changed the results in any way. The fact that these results correspond very well to those obtained elsewhere with a random order (for example Abramson 1962, Liberman et al. 1961), shows that the performance order has at least not had a considerable influence on the results of the tests.

The effect of a change in sound segment duration in separate word positions on the listener's answers is dealt with in what follows and an attempt is made to indicate those factors which have caused a change in the listeners' identifications. The results of each test are shown in the form of a table and of a graphic diagram. The original durations of the sound segments of each test word, the duration of the changed sound segment of the test samples in order of presentation and the percentage distribution of the answers given by the listeners to each sample are shown in the tables 18_{1-23}. On the top of each diagram there is a graph, too, indicating the original durations in milliseconds. The figures for word-initial plosives refer to the interval between the beginning of the plosive's noise burst and the beginning of the following vowel, not the whole duration of the plosive as is normally the case.

4.2. *Shortening of word-initial vowels*

In all of the word pairs chosen in order to examine the effect of shortening of a long first syllable vowel, the vowel appears word-initially. The reason for the choice of a word beginning with a vowel is that with such a word shortening of the original vowel duration can be made simply from the beginning of the word. When a consonant precedes a vowel, no cut can be made at the beginning of the vowel on account of the formant transition. The shortening of an interconsonantal vowel segment by the method used here would imply a steady formant pattern, sound intensity and fundamental frequency during the sound segment to be shortened. If the quality, pitch-pattern, or intensity of voice change during the cut segment, the cutting could easily result in an audible step (Huggins 1968, p. 4–5). When all the words used in the perception test were read by speaker I.S. leaving short pauses between words in the word list, no formant transitions caused by the end of the preceding word occurred at the beginning of a word-initial vowel. The intonation of the test words read in the form of a list was also even, so that the shortening of the word-initial vowel did not

[1] As in most Finnish dialects (see Terho Itkonen 1964, p. 219) laryngalisation (German fester Einsatz) or a glottal stop also can appear in Standard Finnish at the beginning of a word which phonemically begins with a vowel in contrast to the soft initial breath phase which is identified as a word-initial /h/. The laryngal onset of a word beginning with a vowel can also be geminated on the word boundary when it is preceded by a word ending in a morphoneme{X} (X is always realized at the boundary as a doubling of the following consonant): *painu ulos* phonetically [painuˀ: ulos] 'get out'. See chapter 2.1.

affect the pitch-pattern of the word. However, the difficulty in shortening was that the irregular periods characteristic of the beginning of a word-initial vowel and caused by glottalisation were being cut out.[1] When a vowel was shortened in such a way that the beats of the hard onset were cut out and the amplitude of an oscillation increased very quickly, a crack could be heard at the beginning of the vowel, which the listeners could recognize in front of various vowels as various plosives. But when the increase in intensity at the beginning of the vowel abated, some of the listeners recognized the words as having an initial /h/. In the present study the short delay time considered suitable after the preliminary tests did not result in any observations of a word-initial /h/. The same time constant was also used in shortening word-final vowels in order to remove the final crack.

Glottalisation also makes the measurement of the duration of a word-initial vowel uncertain, because the boundary between the glottalisation segment and a pure vowel is often impossible to determine. Since a large part of the measurements of the initial vowel of words beginning with a vowel had to be disregarded because it was impossible to determine precisely the segment boundary, the so-called statistical absence element in this group is so large that words beginning with vowels are excluded from the statistical analyses of the former chapters.[1] Therefore the results of analysis of the durations of the initial vowels of words beginning with vowels read in the sentence frame are only dealt with in this connection.

There were altogether ten pairs of words beginning with vowels in the sentences list. Of the two hundred sentences (20 sentences read by ten speakers) it was possible to define the boundary of a word-initial single vowel in 78 cases, and of a double vowel in 85 cases. According to the measurements, the duration of a word-initial single vowel is on average 69 msec and the duration of a word-initial double vowel on average 162 msec, i.e. 2.3 times as long as the duration of a word-initial single vowel. The average durations of word-initial single and double vowels are almost the same as the durations of single and double vowels in the first syllable of the words beginning with consonants in the list of

[1] The boundary is relatively easy to determine in cases where the final sound of the previous word is a nasal, for example [*melkein 'uudella*], likewise in cases where there is a long laryngal plosive on the word boundary, for example [*koeta ˀ:udella*]. However, glottalisation between a word-final and a word-initial vowel is very often such that the separation from »pure» vowels would be rather arbitrary. The function of glottalisation in Finnish as a boundary signal in word juncture has been studied by Lehiste 1965 B, p. 174–178.

meaningful sentences (70 msec and 153 msec; VV/V ratio 2.2). Hence the duration of a word-initial vowel can be said to correspond to the duration of the vowel of a first syllable beginning with a consonant and one can consequently assume that observations made of the shortening of a word-initial long vowel also hold true for the duration of the vowel of a first syllable which begins with a consonant.

The four words, *uuni* 'oven', *Eetu* (person's name), *Eesti* 'Estonia', *uudella* 'with the new', each of which has a single vowel structural partner *uni* 'sleep', *etu* 'advantage', *esti* '(he) prevented' and *udella* 'inquire' were chosen for a test in which the first vowel of each word was shortened by 15 msec at a time. In the lists read by speaker I.S., from which the test words have been chosen, the average speech tempo is slower than with the ten speakers who read the sentences of the analysis tests. The durations of word-initial single vowels of the above-mentioned words read by I.S. is exactly twice as long, and the duration of double vowels almost 2.4 times as long as the average durations of the corresponding sounds of words read by the ten speakers within a sentence frame in a tertiary sentence stress position.[1]

With the exception of the *udella/uudella* -pair, in which the main difference between the structures is in the duration of the first vowel segment, words with a single initial vowel and double initial vowel also differ in the duration of other sound segments. The second syllable vowel in (C)VCV-words is longer than the first syllable vowel and is clearly longer than the second syllable vowel in (C)VVCV words. In the word *esti* the initial member /s/ of the intervocal consonant cluster is longer in duration than the following member of the cluster and longer than the /s/-segment of the word *Eesti*.

In spite of the differences between double and single vowel words, shortening of the duration of the original double vowel led to its recognizition as the corresponding single vowel word. In tables 1–4, which show the distribution of the answers to the four tests one can, however, see that in all cases the duration of the initial vowel must have been considerably shorter than the original single vowel before the words could be recognized as starting with a single vowel.

In test Nr 1 the stimulus is recognized 100 % as the word *uni*, the duration of the V_1 being 120 msec: the duration of the /u/ in the original word *uni* read

[1] In spite of such great differences one can state that the tempo of both speaker I.S. and the ten sentence list readers is normal.

TABLES 18. $_{1-4}$. Listeners' responses to stimuli prepared by varying the duration of the originally long initial vowel in four test-words.

Test word Nr 1. Orig. *uuni* 'oven'

uu (435) n (85) i (100)

[u] msec	»uni» %	»uuni» %
60	100	–
75	100	–
90	100	–
105	100	–
120	100	–
135	90	10
150	38	62
165	21	79
180	5	95
195	2	98
210	–	100
225	–	100

Test word Nr 2. Orig. *Eetu* (person's name)

ee (360) t (130) u (90)

[e] msec	»etu» %	»eetu» %
40	100	–
55	100	–
70	100	–
85	100	–
100	100	–
115	86	14
130	57	43
145	21	79
160	7	93
175	2	98

Test word Nr 3. Orig. *uudella* 'with the new one'

uu (360) d (45) e (140) ll (190) a (85)

[u] msec	»udella» %	»uudella» %
230	–	100
215	–	100
200	–	100
185	–	100
170	2	98
155	7	93
140	36	64
125	71	29
110	86	14
95	100	–
80	100	–
65	100	–

Test word Nr 4. Orig. *eesti* 'Estonia'

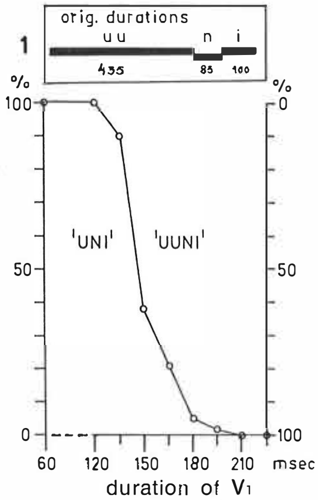
ee (290) s (130) t (130) i (60)

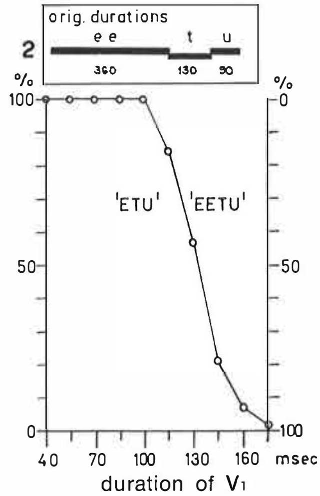
[e] msec	»esti» %	»eesti» %
60	100	–
75	100	–
90	100	–
105	95	5
120	86	14
135	36	64
150	2	98
165	–	100
180	–	100
195	–	100

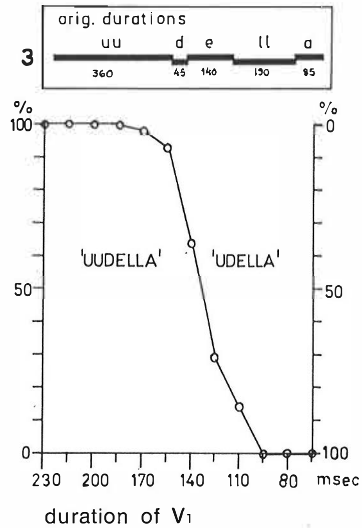
by I.S. in the corresponding word list is 140 msec. The V_1 is accepted as a single vowel in test 2, its duration being 100 msec. The duration of V_1 in the original word *etu* is 140 msec. In test 3 all the listeners recognized the stimulus as the word *udella* with a V_1 duration of 90 msec; the duration of the /u/ in the original *udella* is 130 msec. The greatest durational difference between an original single vowel and a single vowel shortened from a long vowel is in the pair *Eesti*/*esti*. The test word is accepted as single-vowelled only when the duration of the V_1 is 90 msec, even though the duration of the V_1 of the original word *esti* read in the corresponding position is 156 msec.

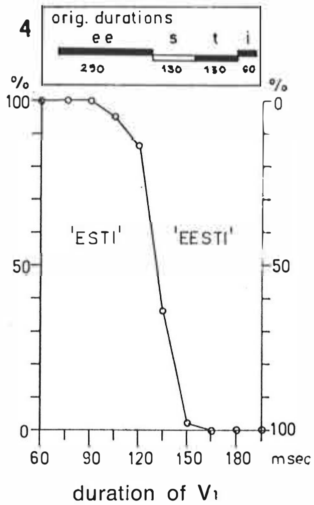
The durational difference between the original single vowel and the shortened long vowel recognized as a single vowel may be related to the differences which were seen to exist between the duration in other sound segments of original double and single vowel words. The recognition of a vowel segment as a single or double phoneme is probably affected by the ratio of the duration of the changed sound segment the duration of other phonemes in the word. In the (C)VCV/(C)VVCV structure pairs (p. 106) only the duration of the V_2-segment showed a clear change besides the duration of the V_1-segment: When in (C)VVCV-word V_1 within a sentence is on average three times as long in duration as V_2 the V_1/V_2 ratio in the (C)VCV-words is reversed: V_2 is on average over 1.5 times as long in duration as V_1. In the test words representing a slow tempo the ratios are slightly different, but have the same tendency. However, the contrast ratio $V_1<V_2$ cannot be a necessary condition of recognition of the structure (C)VCV because all the listeners recognized both the original words *uuni* and *Eetu* as VCV-words though V_1 was still 20 % longer in duration in the former, and slightly over 10 % longer in the latter than V_2. All the listeners recognized a double vowel in test-words 1 and 2 when V_1 was about 2 times the length of V_2 in duration. The foregoing ratios do not hold true in the recognition of the initial vowel of the three syllable words *udella*/*uudella*, because in (V)VCVCCV(C) structures V_2 is in a different durational position from V_2 in *uni*/*uuni* words. The duration of the initial vowel of the word *eesti* had to be made more than one third shorter than the initial vowel of the original word *esti* before all the listeners accepted the vowel as short. The recognition of the vowel might have been affected by the same V_1/V_2 contrast ratio which obviously has an effect on the (C)VCV/(C)VVCV distinction in the recognition of the first vowel. A more important cue affecting the phonemisation of the duration of the V_1 segment could, however, be the durational distribution of the consonant cluster.

Figures 10_{1-29} (pp. 162–184). Listener responses to changed durations of various sound segments of the original words. The durational pattern of the original word is illustrated at the top of each figure. The distribution of the responses is given in numeral form in tables 18_{1-29} and is discussed in details in the corresponding points of the text.









The most considerable difference in the sound segments of the original words in the pair *esti/eesti* – in addition to the initial vowel – is in the duration of the initial member of the consonant cluster.

If the C_x/C_y contrast of the test word as well as the duration of the whole CC-segment, is typical of the (C)VVCCV word, the V_1 segment must be shorter than the V_1 of the original *esti*-word, in order to obtain the corresponding V_1/C_X contrast ratio.

The V/VV -distinction of the first syllable can, according to the tests, be said to be segmental in character in that by shortening a long vowel the corresponding single vowel word can be obtained: in other words the first syllable double vowel can also be commuted by a short vowel in the phonetic sense *ceteris paribus*. The changes in the original duration were made in the tests with the double vowel words. However, it may be that the lengthening of the V_1-segment of the original (C)VCV words would not lead, correspondingly, to recognition of a (C)VVCV-word, because at a certain point the V_1/V_2 contrast ratio would come very near to the corresponding ratio of the (C)VVCVV-structure because of the half-length of V_2 in the (C)VCV structure.

On the basis of test results one can define the smallest durational difference needed to distinguish a first syllable double vowel from a single vowel when the environment remains unchanged. The longest V_1 duration accepted by 100 % of the listeners as a single vowel is on average 101 msec and the average shortest duration recognized by all the listeners as a double vowel is 191 msec. Between these values there is a theoretical phoneme boundary, or a V_1 duration, which would probably be recognized by half of the listeners as a double vowel and by half as single. The phoneme boundary, or critical duration, can be exactly defined in the graphic diagrams representing the distribution of listeners' responses in each test. The phoneme boundary which is, of course, only valid in each individual case, can be defined in the figures by seeing at which duration of the X-axis the line showing the distribution of answers crosses the 50 % level. In the *uuni/uni* -test the crossover point is at 146 msec, in the *eetu/etu* test at 133 msec in the *uudella/udella* test at 134 msec and in the *eesti/esti* test at 130 msec. These center points of the phoneme boundary correspond almost completely to the V_1 duration in the original short vowel words! Recognition of the phonemic quantity of vowels could not thus have taken place by a comparison of the heard vowel duration with the listener's knowledge of the durations of a normal single vowel word appearing in the corresponding (tempo and stress) position, but by a comparison of the contrast ratios between the vowel and the other sound segments of the heard word with typical contrast ratios familiar to the listener appearing in different structures. The fact that the listeners' choice is really made between word structures and not between the quantities of individual sound segments will be even more clearly illustrated in the test results of the other structures in a later section.

4.3. *Shortening of the second syllable vowels*

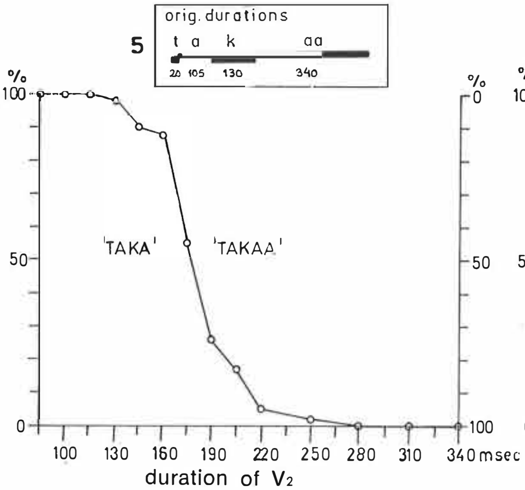
4.3.1. Shortening of the second syllable vowel of the original CVCVV words

In tests 5, 6, 7, 8, 9, the second syllable vowel segment of the CVCVV-type words *takaa*, *mätää*, *pesee*, *likaa* and *mutaa* was shortened in stages of 15 msec from the end. The shortening of final vowel in the words *takaa*, *mätää* and *pesee* simply led to their identification as the corresponding CVCV-wors *taka*, *mätä* *pese*. However, when the final vowels of the original words *mutaa* and *likaa* were shortened, they were also identified apart from the expected *lika*, *muta* partly as the words *liika*, *liikaa* and *mutta*, *muuta*, *muuttaa*. The figure $10._{6}$, which shows the identifications caused by the shortening of the final vowel of the original word *mätää*, contains two tests. The circles joined by an unbroken line indicate the distribution of the listener's answers, when the samples were presented from the longest to the shortest. The triangles joined by a broken line indicate the distiribution of the listener's answers in the test in which the same samples were in the reverse order.

According to the analysis of the durational differences between words read in the sentence frame, the CVCVV/CVCV-distinction is phonetically based on the durational difference between the V_2, and to some extent also between the C_2, segments. The double vowel V_2 is about one and a half times as long in duration as the V_2-segment of the CVCV-structure. It was also shown as a characteristic feature of the CVCV-structure that the second syllable vowel is about 1,5 times the duration of the first syllable vowel in duration. In tests 5–7 the ratio between the longest durations which were wholly (= 100 %) perceived as a single vowel and the shortest duration which were wholly (= 100 %) perceived as a double vowel is on average approximately 1 :1.8. This ratio varies considerably from test to test, but the differences are only apparent and are caused, for example, by »tails» appearing in the *takaa*-test and in the *mätää*-test b), in which a different identification by one listener considerably shifts the 100 % boundaries.

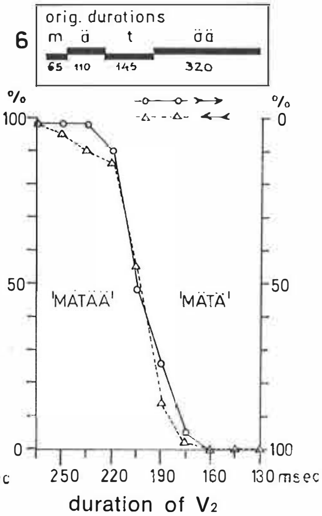
The durations of the initial syllable vowels of the test words can also be compared to that V_2 duration which would probably have caused just as many identifications of single and double vowels. This critical V_2 duration is 1.9 times the duration of V_1 in both *mätää*-tests and in both the *takaa* and the *pesee*-test 1.7 times. The majority of the listeners identified the V_2 of the test words as a double vowel when it was more than twice as long as the V_1

5



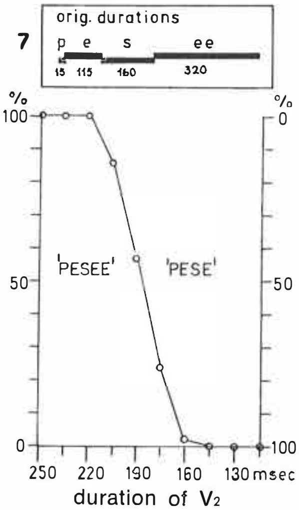
orig. durations
t a k aa
20 105 130 340
%
'TAKA'
'TAKAA'
duration of V2
msec

6



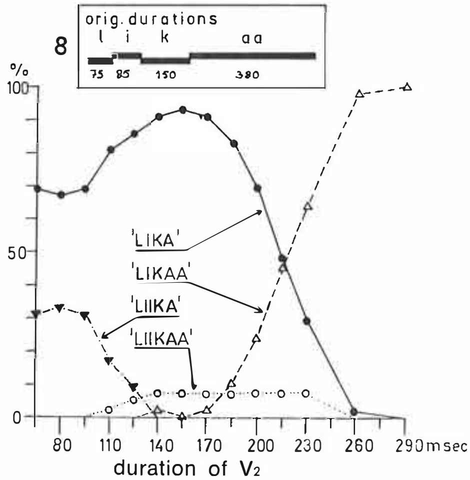
orig. durations
m ä t ää
65 110 145 320
%
'MATÄÄ'
'MÄTÄ'
duration of V2
msec

7



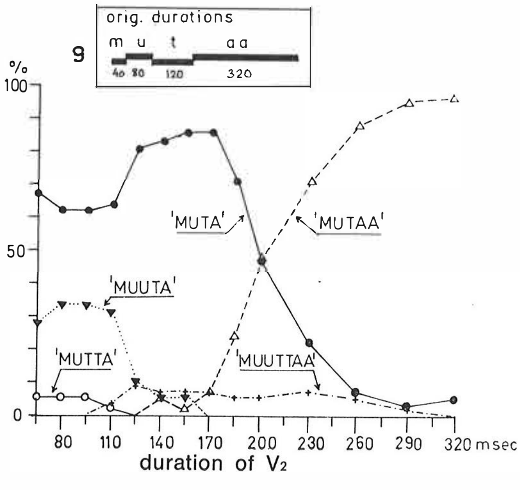
orig. durations
p e s ee
15 115 160 320
%
'PESEE'
'PESE'
duration of V2
msec

8



orig. durations
l i k aa
75 85 150 380
%
'LIKA'
'LIKAA'
'LIIKA'
'LIIKAA'
duration of V2
msec

9



orig. durations
m u t aa
40 80 120 320
%
'MUTA'
'MUTAA'
'MUUTA'
'MUTTA'
'MUUTTAA'
duration of V2
msec

in duration. When the V_2 was 1.5 times the duration of the V_1 the majority identified it as a single vowel in all the four tests.

The listener responses of tests 8 and 9 differ clearly from the former ones. The V_2 duration which caused just as many identifications of *lika* and *likaa* (about 215 msec) – resp. *muta* and *mutaa* – (200 msec) corresponds here to the V_1/V_2 contrast ratio 1 : 2.7 and 1 : 2.8. This exceptional contrast ratio is partly caused by the fact that in test words 8 and 9 the first vowel is an intrinsically short close vowel and the second vowel is an intrinsically longer vowel, but also that the crossover-point of the single vowe- and double vowel answers is on average about 20 msec longer in tests 8 and 9 than in tests 5–7. The appearance of CVVCV identifications *liika* and *muuta* which do not appear in the answers of tests 5–7 is unexpected in tests 8 and 9 as the duration of the V_2 varies between 65 msec – 125 msec. This may be the result of several different factors: the shortest V_2 duration appearing in the stimuli of tests 5–7 is 85 msec and even that is only in the *takaa*-test. However, the V_2 duration of the first stimulus in tests 8 and 9 is only 65 msec. The duration pattern heard first in the test did not correspond to the normal distributions of any bisyllabic structure. If it is supposed that the listeners tend to look for a meaningful interpretation for this unusual stimulus, several possibilities exist. Clearly when the duration of the intervocal consonant is so short that it cannot be formed as a double consonant, the choice is between two structures: the CVVCV and CVCV. Even though the expected vowel-segment durational contrast $V_1<V_2$ of the latter did not appear the majority of listeners accepted this structure. Some of the listeners, however, considered the lack of the typical CVCV durational contrast as a decisive cue and chose the CVVCV identification. However, CVVCV recognition was obviously made possible only by the phonetical similarity between the initial consonant and the following vowel. In the word *mutaa* the grave /u/ -vowel is preceded by a nasal similar to it in colour and in the word *likaa* the /i/ vowel is preceded by a clear allophone of /l/. In the identification of the stimulus which was ambivalent in its durational pattern, it was possible, because of the smallness of the CV-contrast, to perceive the vowel segment as longer by compensating the internal durational relations of the C- and V-segment of the syllable.

About one third of the listeners made the CVVCV-interpretation in the first stimuli of tests 8 and 9, but when the duration of V_2 increased in the following stimuli, these listeners also formed the word again as the expected CVCV-word. Only less than 10 % of the listeners considered the first vowel long even in their later answers where a word representing the CVVC(C)VV structure replaced the CVVCV-word.

The absence of CVVCV words in tests 5–7, may be caused by the fact that the durational pattern was not ambivalent here to such an extent as in tests 8 and, 9 but also that CVVCV identification would not have been lexically meaningful here, because the phonemically possible words **taaka*, **määtä* and **peese* do not occur in Finnish. When a stimulus is phonetically ambiguous the listener automatically chooses the meaningful phonemisation in each context. When the test situation did not offer redundant information of either the speech situation or sentence environment, the lexically and morphologically meaningful interpretation was chosen in place of the nonsense-phonemisation in an ambivalent case. However, the answers to tests 22 and 25 show that even in the language as a word or a morpheme nonexistent phonemisation is chosen, if the pressure of the phonetic cues is great enough.

In the analysis of the sound durations of words appearing in a natural sentence environment, it was shown that in the CVCV/CVCVV distinction the duration contrast between structures is considerably smaller than in the other bisyllabic structure pairs. However, these tests show that listeners are also normally able to distinguish between these structures just on the basis of the durational difference between the V_2-segments.

4.3.2. Shortening of the original second syllable vowel after a geminate consonant

In tests 10, 11 and 12 the durationof V_2 of two original CVCCVV-words (*takkaa* and *kassaa*) and of one CVVCCVV-word (*taakkaa*) with a double vowel following a double consonant was altered in 15 msec stages. In all these tests the shortening of the duration of the V_2 led to its identification as the corresponding single vowel. Those stimuli where the duration of V_2 is long are accepted as the original test-words, with the exception of test 12 where a couple of listeners recognized the stimuli with a relatively long V_2 duration as the CVCV word *taka*. The mean ratio of the longest duration accepted as single to the shortest V_2 duration accepted as a double vowel in the three tests equals 1:2. According to the tests, the second syllable vowel seems to be recognized in the CVCCV(V) words as double when it is two times the length of the V_1 in duration and as single when it is as long as the first vowel. In the CVVCCV(V) words, on the other hand, V_2 is accepted as a single vowel when its duration is about half of the duration of the first syllable double vowel.

The longest duration of V_2 shortened from the original double vowel which was accepted as the single vowel is longer in all three tests than the duration of the single V_2 of the original *takka*, *kassa* and *taakka* words read by I.S. in the corresponding sentence positions. The difference is on average 35 msec (46 % fo the duration of the original single vowel). The V_1/V_2 contrast ratio of the stimuli in tests 10 and 11 accepted as CVCCV words does not correspond to the corresponding ratio of the original CVCCV words either: while the V_1/V_2 of the CVCCV words read in an unstressed position within a sentence is on average 1 : 0.6, the corresponding V_1 ratio to the longest accepted single duration vowel of V_2 in test 10 is approximately 1 : 1.2 and in test 11 exactly 1 : 1. The corresponding contrast ratio between first and second syllable vowels falls

TABLES $18._{5-9}$ Listener responses to stimuli prepared by shortening the originally double word final vowel segment preceded by a single consonant

Test word Nr 5. Orig. *takaa*
t (20) a (105) k (130) aa (340)

[a_2] msec	»taka» %	»takaa» %
85	100	–
100	100	–
115	100	–
130	98	2
145	90	10
160	88	12
175	55	45
190	26	74
205	17	83
220	5	95
250	2	98
280	–	100
340	–	100

Test word Nr 6. Orig. *mätää*
m (65) ä (110) t (145) ää (320)

[$ä_2$]	»mätä»		»mätää«	
msec	I	II	I	II
130	100	100		
145	100	100		
160	100	100		
175	95	98	5	2
190	74	86	26	14
205	52	45	48	55
220	10	14	90	86
235	2	10	98	90
250	2	5	98	95
265	2	2	98	98

Test word Nr 7. Orig. *pesee*
p (15) e (115) s (160 ee (320)

$[e_2]$ msec	»pese« %	»pesee« %
250	100	–
235	100	–
220	100	–
205	86	14
190	57	43
175	24	76
160	2	98
145	2	98
130	–	100
115	–	100

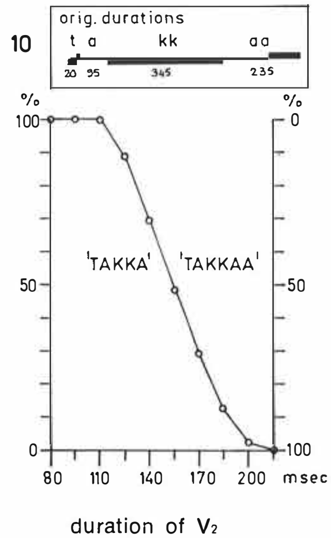
Test word Nr 8. Orig. *likaa*
l(75) i (85) k (150) aa (380)

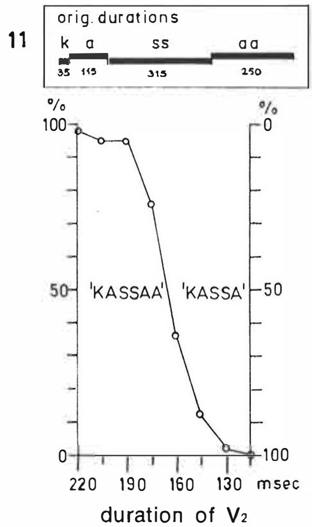
$[a_2]$ msec	»lika« %	»likaa« %	»liika« %	»liikaa« %
65	69	–	31	–
80	67	–	33	–
95	69	–	31	–
110	81	–	17	2
125	86	–	9	5
140	91	2	–	7
155	93	–	–	7
170	91	2	–	7
185	83	10	–	7
200	69	24	–	7
215	48	45	–	7
230	29	64	–	7
260	2	98	–	–
290	–	100	–	–

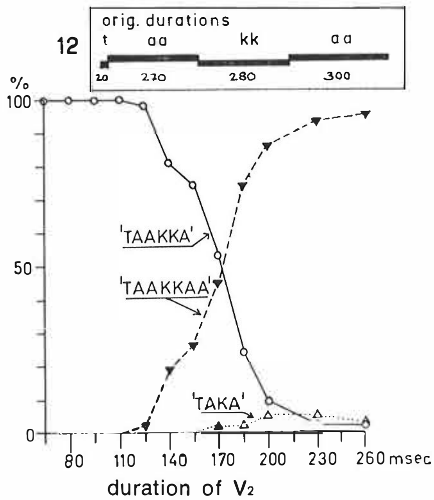
Test word Nr 9. Orig. *mutaa* m(40) u (80) t (120) aa (320)

$[a_2]$ msec	»muta« %	»mutaa« %	»muuta« %	»muuttaa« %	»mutta« %
65	67		28		5
80	62		33		5
95	62		33		5
110	64		31	3	2
125	81	–	10	9	–
140	83	5	5	7	–
155	86	2	5	7	–
170	86	7	–	7	–
185	71	24	–	5	–
200	47	48	–	5	–
230	22	71	–	7	–
260	7	88	–	5	–
290	3	95	–	2	–
320	5	95	–	–	

between those (1 : 0.8) on the original CVCCV words read in a list by I.S. However, the contrast ratio of the test corresponds approximately to the V_1/V_2 contrast ratio of the CVCCV words read very slowly by ten readers, where the durational difference between the first and second syllable single vowels almost entirely disappeared.







The results of the tests seem to indicate that identification of the stimuli took place here according to the same patterns as the production of structures in slow speech. In the same way as the durational difference between various vowels and the different effects of various consonants, the different effects of various word positions on the duration of sound segments are also levelled out in slow speech and identification of the structures takes place according to different patterns from those in quick speech. Intervocal double consonants, which are lengthened about 100 % compared to those within sentences when the lengthening of 1st syllable single vowels is only 30 %, act as indicators of slow tempo in the test stimuli. Therefore the recognition of the V_2 does not take place on the basis of the V_1/V_2 contrast but rather on that of the C_2/V_2 contrast.

4.4. *The changes in the duration of intervocal consonants*

In tests 13–29 the duration of the intervocal consonants has been changed: in tests 13–19 the objects of the changes are the original CVCV words, in tests 20–25 the CVCCV words and in tests 26–29 the words representing the CVCCVV structure. In all test words either plosives /ptk/ or the fricative /s/ stand as the intervocal consonants. In the method used here the plosive

TABLES $18._{10-12}$. Listener responses to stimuli prepared by shortering the originally double word final vowel preceded by a double consonant

Test word Nr 10. Orig. *takkaa*
t (20) a (95) kk (354) aa (235)

$[a_2]$ msec	»takka» %	»takkaa» %
80	100	–
95	100	–
110	100	–
125	88	12
140	69	31
155	48	52
170	29	71
185	12	88
200	2	98
215	–	100

Test word Nr 11. Orig. *kassaa*
k (35) a (115) ss (315) aa (250)

$[a_2]$ msec	»kassa» %	»kassaa» %
220	2	98
205	5	95
190	5	95
175	24	76
160	64	36
145	12	88
130	2	98
115	–	100

Test word Nr 12. Orig. *taakkaa*

(20) aa (270) kk (280) aa (300)

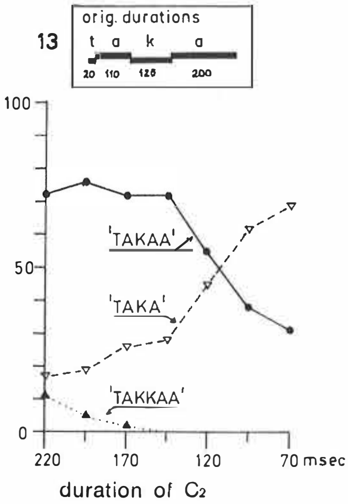
$[a_2]$ msec	»taakka« %	»taakkaa« %	»taka« %
65	100	–	–
80	100	–	–
95	100	–	–
110	100	–	–
125	98	2	–
140	81	19	–
155	74	26	–
170	53	45	2
185	24	73	2
200	9	86	5
230	2	93	5
260	2	95	3

segment gap and the steady state of the fricative noise-segment could be lengthened to only twice their original duration. Care was taken in the choice of the fricative test-words so that the noise-segment of the fricative was as even as possible in quality and intensity. The result of the shortening or doubling of the fricative segment was quite equal to the result of changes in the voiceless part of plosives: the cutting was in no way audible although it is instrumentally observable in the fricative segment. The imperceptibility of the cutting is also indicated by the fact that when the listener hears the words, one of which is changed in the duration of its intervocal consonant, he is often only able to recognize them as different in their durational patterns without being able to ascertain which sound duration has changed.

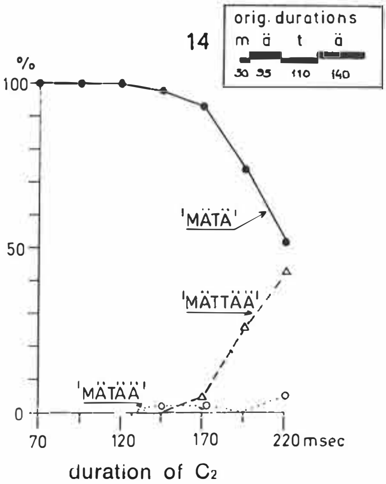
4.4.1. The changes in the duration of the intervocal single consonant

In tests 13–19 the effect of changes in the duration of the intervocal consonant segment in originally CVCV words on the recognition of the words was examined. It was possible to lengthen the duration of the consonant segment only until it was double the original duration, which in the test was only partly sufficient for the identification of a double consonant.

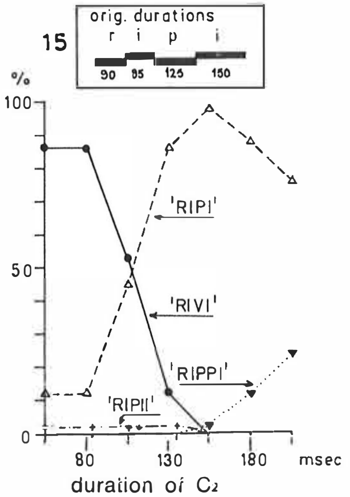
13



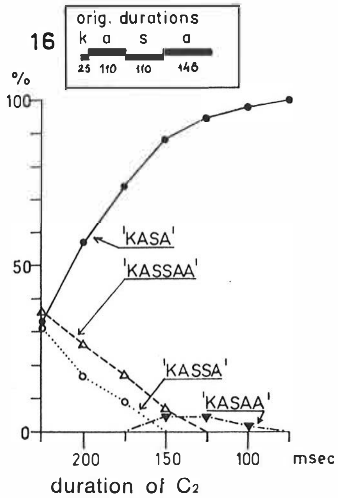
14



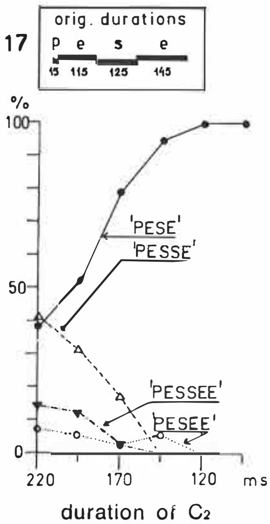
15



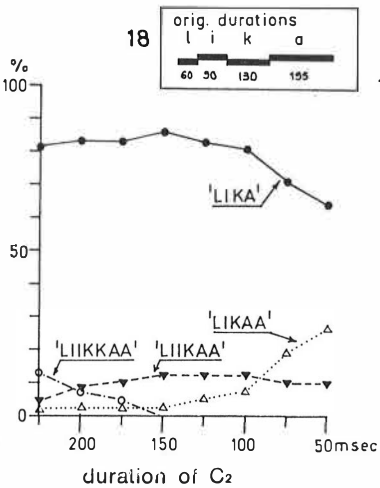
16



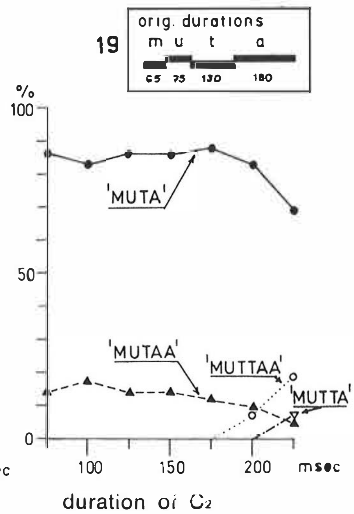
17



18



19



In natural speech the duration of an intervocal single consonant is generally influenced by the length of the following vowel. The consonant is always phonetically longer in duration before a long vowel than before a short vowel. Since the durational difference seems fairly regular according to the analysis, one could assume that it would also be used in perception as a recognition cue for the phonemic length of the vowel following the consonant. In particular, this could be expected in the CVCV/CVCVV distinction where the durational difference between the second syllable single and double vowels is clearly less than in other positions. The possibility that the consonant might act as a recognition cue for the following vowel is also an interesting subject for observation because the several dialects of Finnish sharpened the original CVCV/ CVCVV distinction by lengthening a consonant before a long vowel element, i.e. by so-called general gemination, and thus arrived at the structural contrast CVCV:CVCCVV (for example /kala/ 'fish': /kallaa/ 'of the fish' part.sg.).

If consonantal duration is also used in the standard language under consideration as a recognition cue for the following vowel, it could be expected that the consonant lengthening would cause observations of the CVCVV word at least in cases where the second syllable half-long vowel is longest. The distribution of the answers of test 13 would seem to indicate that the duration of a consonant really affects evaluations of the length of the following vowel. The lengthened duration of C_2 caused over 70 % *takaa* recognitions. This percentage decreases as the duration of *k* gets shorter. In tests 14–17 the amount of the CVCVV identifications is, however, small and it is not clearly related to the increase in the duration of the intervocal consonant. On the other hand in the answers to tests 18 and 19 the CVCVV identifications are related, in contrast to test 13, to the short durations of the intervocal consonant. As a whole the results of tests 13–19 seem to indicate that the duration of a single consonant following a short syllable is not systematically used as an identification cue of the length of the vowel following a consonant. The obvious reason for the exceptional answers in test 13 is in the original which is already ambiguous in phonetical duration. While in all the other tests of this series identification of the stimulus corresponding to the original duration of the C_2 is over 80 % correct for the CVCV word, more than half the listeners recognized in test 13 the stimulus which most closely corresponds to the original *k* (duration of *k* 120 msec) as a CVCVV word. In test words 14–16, in which the same vowel appears in positions V_1 and V_2, the V_1/V_2 durational ratio is on average 0,7 (1 : 1.4), whereas the corresponding ratio in the word *taka* is 0.55; in other words a second syllable half-long vowel is about two times as long as a first syllable vowel. However, V_1/V_2 in test words 18 and 19 is even longer (0.45 and 0.42), but the distribution of the answers differs totally, nevertheless, from the distribution in the *taka* test. The V_1/V_2 ratio in these is, however, affected by the fact that the durationally shorter close vowel is the vowel of the first syllable and the durationally longer open vowel is the vowel of the second syllable. Here, as in tests 8 and 9, the similarity between the initial consonant and the first vowel of a word may have an effect on the identification of stimuli which are unusual in durational distribution.

The test word *ripi* was not ideal in the perception test because the lexical word is fairly rare and partly because in the chosen expression the burst of the plosive may have been exceptionally soft.[1] Appromixately 80 % of the listeners identified the shorter /p/ durations as a homorganic semi-vowel and the word as *rivi* 'row'. Even the original duration of /p/ caused 20 % identifications as *rivi*. The lengthened duration of the intervocal consonant caused, in all the tests, recognitions of a double vowel whose share of the identifications increases steeply at the end of each figure corresponding to longer consonant duration. Besides the change of a consonant to a geminate, there also appeared to be, in the observations, a change of the V_2 to a double vowel in all the other tests except test 15, in which the absence of CVCCVV answers may be explained by the non-existence of the word **rippii* in Finnish. Accordingly, the absence of a CVCCV word in the answers to test 14 can be explained in the same manner: in the interpretation of a possibly ambiguous stimulus a lexically meaningful alternative was chosen. In the answers to test 13 there is also no CVCCV structure word in spite of the existence of the word *takka*. The reason for this may be found in the durational differences between the vowels which may also have caused deviating distribution of the other answers to test 13. The duration of the original V_2 of *taka* has also facilitated CVCVV identification even during the increase in the consonant, but in the others

[1]The word *ripi* 'birch' belongs, above all, to nursery-words and its long vocal equivalent *ripii* 'it is sprinkling' is also rather rare (occasional or dialect) compared with the more common *ripsii, ripsuttelee*. Identification of the shortened duration of /p/ as /v/ was completely in accordance with expectations and the same phenomenon which has affected observations here can be seen to have caused the /p/–/v/ variation in Finnish consonant grading e.g. *napa* 'navel': *navan* 'of the navel' (gen sg.).

lengthening of the consonant caused a durational pattern which was obviously not typical of any of the word structures.

The ambiguity of the stimulus appears in the answers to test 16 which are equally distributed between the structures CVCV, CVCCV, and CVCCVV when the duration of C_2 is 225 msec. In so far as test words 13, 18 and 19 can be considered comparable to words 14–17 the test results seem to indicate that identification of an intervocal consonant depends among other things on the duration of the following vowel: the longer V_2 is in duration the longer the consonant is which is accepted as a single vowel.

TABLES 18. $_{13-19}$. Listener responses to stimuli prepared by varying the duration of the originally single medial consonant segment between single vowels

Test word Nr 13. Orig. *taka*
t (20) a (110) k (125) a (200)

[k] msec	»taka» %	»takaa» %	»takkaa» %
220	17	72	11
195	19	76	5
170	26	72	2
145	28	72	–
120	45	55	–
95	62	38	–
70	69	31	–

Test word Nr 14. Orig. *mätä*
m (30) ä (95) t (110) ä (140)

[t] msec	»mätä» %	»mätää» %	»mättää» %
70	100	–	–
95	100	–	–
120	100	–	–
145	98	2	–
170	93	2	5
195	74	–	26
220	52	5	43

Test word Nr 15. Orig. *ripi*
r (90) i (95) p (125) i (150)

[p] msec	»rivi» %	»ripi» %	»ripii» %	»rippi» %
55	86	12	2	–
80	86	12	2	–
105	53	45	2	–
130	12	86	2	–
155	–	98	–	2
180	–	88	–	12
205	–	76	–	24

Test word Nr 16. Orig. *kasa*
k (25) a (100) s (110) a (145)

[s] msec	»kasa» %	»kasaa» %	»kassa» %	»kassaa» %
225	33	–	31	36
200	57	–	17	26
175	74	–	9	17
150	88	5	–	7
125	95	5	–	–
100	98	2	–	–
75	100	–	–	–

Test word Nr 17. Orig. *pese*

p (15) e (115) s (125) e (145)

[s]	»pese»	»pesee»	»pesse»	»pessee»
msec	%	%	%	%
220	38	7	41	14
195	52	5	31	12
170	79	2	17	2
145	95	5	–	–
120	100	–	–	–
95	100	–	–	–

Test word Nr 18. Orig. *lika*

l (60) i (90) k (130) a (195)

[k]	»lika»	»likaa»	»liikaa»	»liikkaa»
msec	%	%	%	%
225	81	2	4	13
200	83	2	8	7
175	83	2	10	5
150	86	2	12	–
125	83	5	12	–
100	81	7	12	–
75	71	19	10	–
50	64	26	10	–

Test word Nr 19. Orig. *muta*

m (65) u (75) t (130) a (180)

[t]	»muta»	»mutaa»	»muuttaa»	»mutta»
msec	%		%	%
75	86	14	–	–
100	83	17	–	–
125	86	14	–	–
150	86	14	–	–
175	88	12	–	–
200	83	10	7	–
225	69	5	19	7

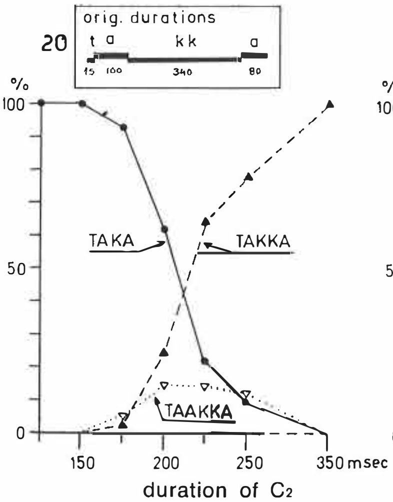
4.4.2. Durational changes in the intervocal consonant of the original CVCCV words

The duration of the intervocal consonant segment of the original CVCCV -structure word was changed in tests 20–25. Compared to the words read at a normal tempo within the sentence frame there are certain changes in the timing of the test words read in the list by I. S. characteristic of slower tempo. In the word structure under consideration the duration of the phonemically double C_2 segment increased by about 100 %, the duration of the single V_2 by about 70 %, but the single V_1 by about only 30 %. Because of the different lengthening of different segments the average contrast ratios in the structure also change along with speech tempo. While V_2 is on average only 60 % of the duration of V_1 in CVCCV words in a sentence-medial position, in the words read by I. S. it is 70–90 % of the duration of the V_1 and in the words read at a consciously slow tempo by the ten readers it was even over 100 %.

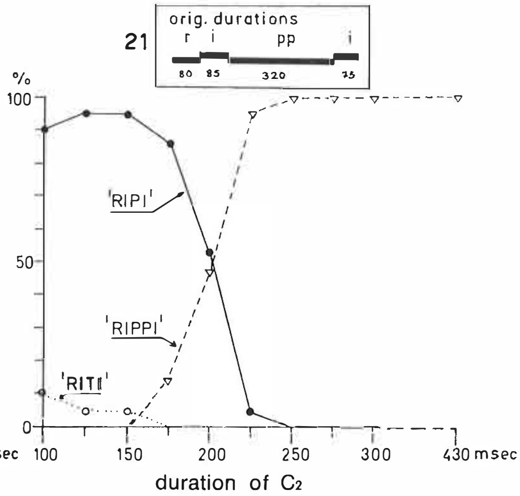
Regardless of test 21, there are words in the answers to all the tests, after the shortening of the consonants duration from the original, in which the first syllable vowel is a double vowel and the intervocal consonant is either a single or double consonant. The fact that the observation of a double vowel is particularly related to the durational contrast of vowels is shown by the answers to tests 23 and 24. When the duration of a consonant was shortened, the number of CVVCV answers increased in both tests up to a C_2 duration of 75 msec at which point it reached its maximum. The last stimulus of the test was so designed that the first vowel of the original words *likka* and *mutta* was shortened by 15 msec and the duration of the consonant of the former stimulus was lengthened to such an extent that it corresponded to the duration of the C_2 of the original words *lika*, *muta* read by I. S. The number of double vowel identifications dropped in test 23 to 10 % and in test 24 to 19 %. Again, regardless of test 21, the distribution of CVVCCV observations is common to all the tests: the identification peak is in the middle of the area of change of consonant duration, near the point of intersection of the lines showing the distribution of the CVCV and CVCCV answers and they decrease as they go in each direction. The cues indicating the CVVCCV structure must have been particularly strong, because the listeners wrote on their forms – in contrast to other tests – words which are phonemically possible but non-existent in Finnish such as **ryyssä*, **kaassa* (in test 22 also **kaasa*) and the descriptive word *liikka* which usually only appears as a component of compound words (for example *liikkakinttu* as compared to the verb *liikata* 'to limp'). The change of an original single vowel to a double vowel in perception and the preservation, however, of a consonant as a double consonant when the duration of the consonant is shortened to no more than half the original can only be explained if the listener's concept of speech tempo is assumed to have changed.[1] Some of the listeners could no longer accept the duration pattern containing a shortened C_2 as a CVCCV word, because the consonant was too short for a geminate,

[1] The influence of changed segment-duration upon the impression of the speech tempo of the whole utterance is also discussed by Huggins 1968, p. 11.

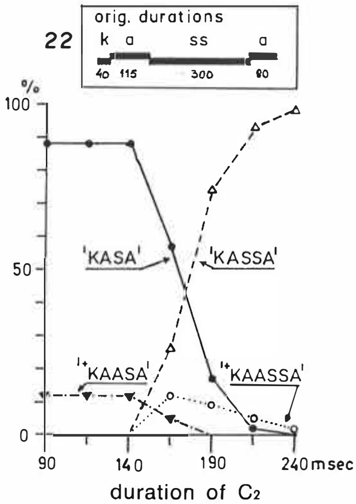
20



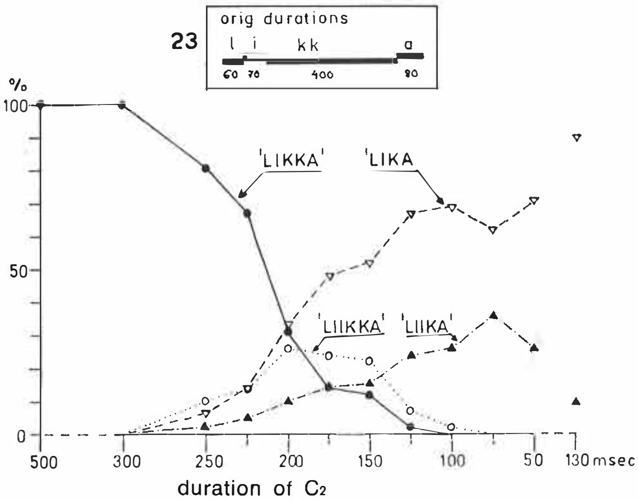
21



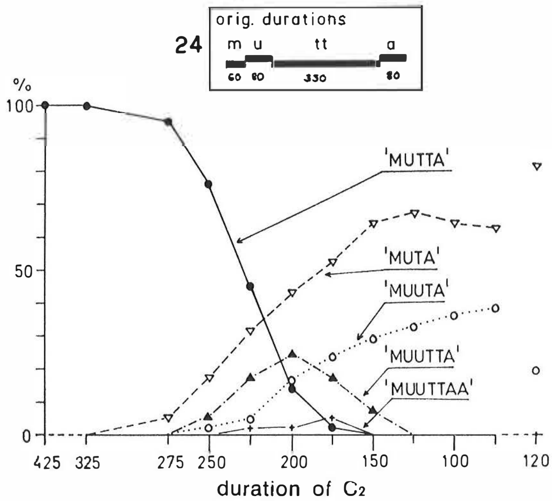
22



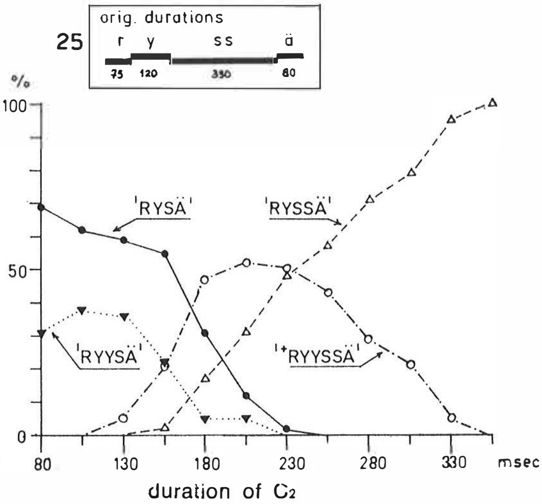
23



24



25



nor as a CVCV word, perhaps because the contrast ratio between the vowels was not typical of a CVCV word. By choosing a »comparator» or perception reference pattern in accordance with the faster speech tempo, the listener could accept the consonant as a geminate, but the V_1 had also to be identified as a double vowel, perhaps because it did not satisfy the $V_1 < V_2$ demand of a CVCV structure of normal tempo. However, the second syllable vowel was not identified as a double vowel except in some of the scattered answers to test 24. One possible explanation is that when V_1 was accepted as a double vowel it was exceptionally short compared to the duration of the consonant; compared to the duration of the vowel the consonant was exceptionally long and according to the C_2/V_2 ratio the duration of V_2 was not sufficient for its identification as a double vowel. However, in all the tests, with the exception of test 25, in which V_1 is exceptionally long in duration, the overall majority of listeners chose between CVCCV and CVCV words, regardless of the unusual V_1/V_2 ratio in the CVCV word. Speakers of standard Finnish have learned to recognize two durational patterns for a CVCV word: the patterns of those who use the so-called half-long vowel and the patterns of those in whose speech the first and second syllable vowel are of approximately the same duration. In accordance with the second of these two comparison patterns, it is possible to identify the stimuli in the tests as CVCV words.

TABLES $18._{20-25}$. Listener responses to stimuli prepared by varying the duration of the originally double consonant segment preceding a single vowel

Test word Nr 20. Orig. *takka*
(80) i (85) pp (320) i (75)

[k] msec	»taka» %	»takka» %	»taakka» %
125	100	–	–
150	100	–	–
175	93	2	5
200	62	24	14
225	22	64	14
250	10	78	12
350	–	–	100

Test word Nr 21. Orig. *rippi*
t (15) a (100) kk (340) a (80)

[p] msec	»ripi» %	»rippi» %	»riti» %
100	90	–	10
125	95	–	5
150	95	–	5
175	86	14	–
200	53	47	–
225	5	95	–
250	–	100	–
–			
430	–	100	–

Test word Nr 22. Orig. *kassa*

k (40) a (115) ss (300) a (80)

[s]	»kassa»	»*kaassa»»	*kaasa»	»kasa»
msec	%	%	%	%
90	–	–	12	88
115	–	–	12	88
140	–	–	12	88
165	26	12	5	57
190	74	9	–	17
215	93	5	–	2
240	98	2	–	–

Test word Nr 23. Orig. *likka*

l (60) i (70) kk (400) a (80)

[k]	»likka»	»liikka»	»liika»	»lika»
msec				
500	100			
– –				
300	100			
250	81	10	2	7
225	67	14	5	14
200	31	26	10	33
175	14	24	14	48
150	12	22	14	52
125	2	7	24	67
100	–	2	26	72
75	–	–	36	64
50	–	–	26	74
– –				
130*			10	90

*) i=55 msec

Test word Nr 24. Orig. *mutta* m (60) u (80) tt (330) a (80)

[t]	»mutta»	»muutta»	»muuttaa»	»muuta»	»muta»
msec	%	%	%	%	
425	100	–	–	–	–
325	100	–	–	–	–
275	95				5
250	76	5	–	2	17
225	45	17	2	5	31
200	14	24	2	17	43
175	2	17	5	24	52
150	–	7	–	29	64
125	–	–	–	33	67
100	–	–	–	36	64
75	–	–	–	38	62
– –					
120*	–	–	–	19	81

*) u=65 msec

Test word Nr 25. Orig. *ryssä*

r (75) y (120) ss (330) ä (80)

[s] msec	»rysä» %	»ryssä» %	»*ryyssä» %	»ryysä» %
80	69			31
105	62			38
130	59		5	36
155	55	2	21	22
180	31	17	47	5
205	12	31	52	5
230	2	48	50	–
280	–	57	43	–
255	–	71	29	
305	–	79	21	
330	–	95	5	
355	–	100	–	

TABLES 18.$_{26–29}$. Listener responses to stimuli prepared by varying the duration of the originally double word-medial consonant segment preceding a double vowel

Test word Nr 26. Orig. *takkaa*

t (20) a (95) kk (345) aa (235)

[k] msec	»taka» %	»takaa» %	»takkaa» %	»taakkaa» %
45	2	98	–	–
70	2	98	–	–
95	–	100	–	–
120	–	100	–	–
145	–	100	–	–
170	2	95	3	–
195	10	76	14	–
220	12	57	19	2
245	12	24	62	2
270	7	10	81	2
295	2	3	95	–
395	–	–	100	–

Test word Nr. 27. Orig. *mättää*

m (30) ä (90) tt (420) ää (250)

[t] msec	»mätä» %	»mätää» %	»mättää» %
155	2	98	–
180	2	98	–
205	2	95	3
230	5	83	12
255	5	57	38
280	5	17	78
305	5	–	95
330	–	–	100
--			
430	–	–	100
--			
590	–	–	100

Test word Nr. 28 Orig. *ryssää*

r (100) y (110) ss (320) ää (265)

[s] msec	»rysä» %	»rysää» %	»ryssää» %	»ryssä» %
50	7	93	–	–
75	7	93	–	–
100	5	95	–	–
125	10	90	–	–
150	12	88	–	–
175	12	83	5	–
200	14	62	22	2
225	17	38	43	2
250	10	19	66	5
275	5	–	90	5
300	3	–	95	2
325	–	–	95	5

Test word Nr. 29 Orig. *liikkaa*

l (50) ii (200) kk (290) aa (210)

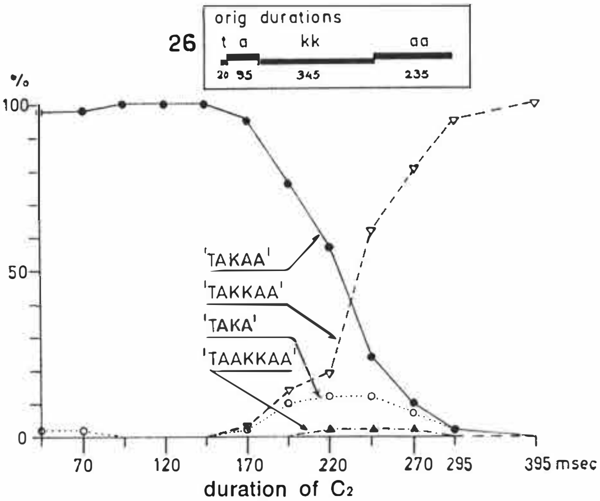
[k] msec	»liikaa» %	»liikkaa» %
365	7	93
315	7	93
290	10	90
265	19	81
240	29	71
215	67	33
190	93	7
165	100	–
– –		
90	100	–
65	93*	

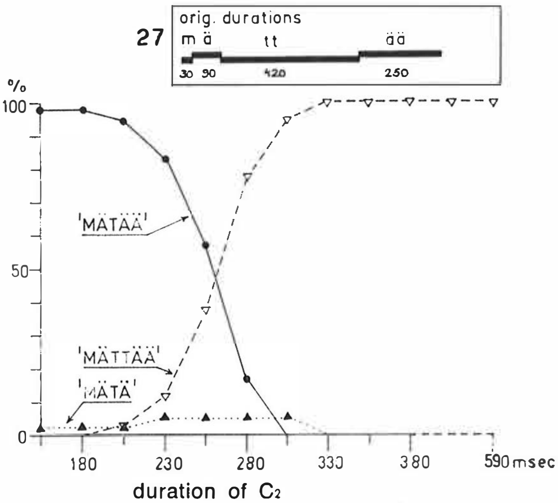
*) 2 responses »liiga», 1 »lika»

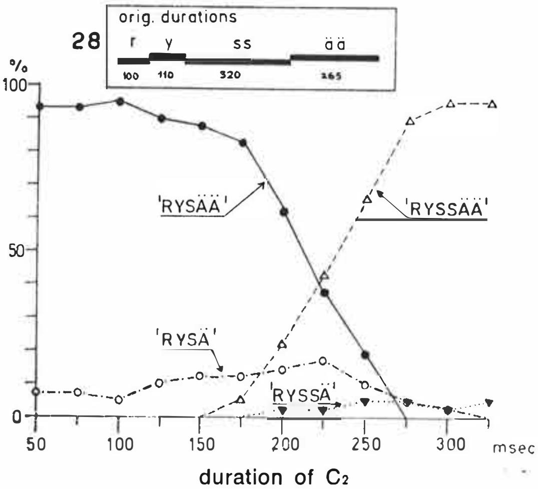
4.4.3. Shortening of an original double consonant preceding a second syllable double vowel

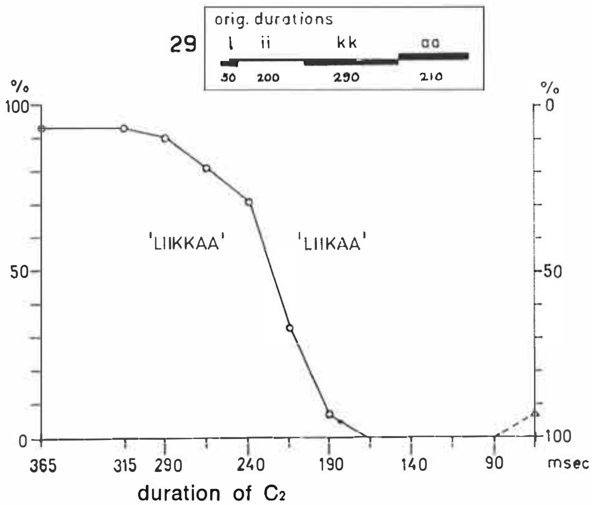
In tests 26–29 the duration of the intervocal consonant in the three CVCCVV words (*mättää*, *takkaa* and *ryssää*) and one CVVCCVV word (*liikkaa*) was altered. In the latter case shortening of the consonant merely led to its identification as the corresponding CVVCVV word *liikaa*. The shortest duration accepted as a double consonant in test 29 is almost twice as long as the duration of *k* wholly (100 %) accepted as a single consonant.

In each of the three CVCCVV tests there is a small group of CVCV answers: in fact, in the *mättää*-test they only represent, at the utmost, 5 % of all the answers. In both the *takkaa* and the *ryssää* test the maximum number of CVCV answers occurs at the point of intersection of the CVCVV and CVCCVV answers. This factor is a further indication, therefore, of the significance of the C_2/V_2 ratio in the perception of these structures: when C_2 is long in duration but not long enough from the point of view of the C_2/V_2 ratio to be identified as a geminate, $/VV/_2$ may be identified as a 'half-long' single vowel in the









CVCV structure regardless of the great V_1/V_2 contrast. It is interesting to note that a partial »overlapping» occurs even here, exactly between the CVCV and CVCVV structures.

However, the majority of identifications in all the tests is of a CVCVV word when the duration of the consonant is shortened. If it is assumed that the speech tempo has remained the same in all the word lists read by I. S. it is possible to compare the results of consonant shortening in the CVCCV words in tests 20–25 and in the CVCCVV words in tests 26–28. The point at which the lines indicating CVCCV and CVCV answers intersect in tests 20–25 corresponds to an average consonant duration of 200 msec. The duration of the consonant in tests 27–28 which caused just as many CVCCVV and CVCVV identifications is on average 242 msec, and the corresponding crossover-point in test 29 is at the 225 msec duration point of the plosive. Therefore, in the compared tests an intervocal consonant segment is accepted as a single consonant before a long vowel when it is on average 40 msec longer than before a short vowel. The value of the C_2/V_2 ratio which contributes to the phonemisation of the duration of C_2 depends on the intrinsic duration of the consonant. In the test-words where the consonants in C_2 position were the obstruents /tks/, the intervocal consonant was identified as single if it was shorter than the following double vowel in duration, and as double if it was longer than the following vowel in duration. The stimulus was ambiguous, when the duration of the intervocal consonant was the same as that of the following vowel. These ratios, however, are not valid in connection with consonants which are short in intrinsic duration.

The number of other answers in tests 26–28 is small that no general conclusion can be drawn from them. (»taakka» in test 26, »ryssä» in test 29 and »liigaa» in test 29).

4.5. *Conclusions and assessments*

The identification tests made by means of artificially changed segment durations show that the identification of a single sound segment as a single phoneme or as a combination of two similar phonemes depends on a realization of the timing relations between the sound segments of the whole linguistic signal, which also give cues to the listener concerning the speech tempo and dialectal or individual reference pattern by which he can decode the speech signal. In the recognition of an individual word the most important cues were shown to be, not the absolute durations of sounds, but the contrast ratios between the durations of the sound segments of a word: the intervocal duration ratio between the first and second (stressed and the following unstressed) syllable and above all the ratio between the duration of an intervocal consonant and the preceding and following vowel.

The analysis and identification tests may also provide a reason for assessing the significance of the timing contrasts between segments in the automatic production of Finnish and even in automatic identification. If in automatic production only understanding, but not necessarily naturalness, is put forward as a timing condition, the timing patterns of different word structures can, at least as far as the vowel segments are concerned, be made by using the simple durational ratio 1:2 between single and double sounds in all positions. The corresponding ratio is possible even with consonants, at least with the consonants /p t k s/ which are long in intrinsic duration. However, the effect of varying duration on the identification of liquid and nasal segments which are shorter in intrinsic duration is to be explained only by means of synthetic speech.

If one moves in the automatic production of speech to the more detailed timing of segments timing rules become more complicated as the fidelity of natural speech becomes higher. The value of the original segment duration must be modified, not only by the coefficient of tempo and sentence-stress position, but also by the constants of intrinsic sound duration, intrinsic effect of adjacent sounds and word position.[1] In a rough simulation three combinatory duration variants might be sufficient to represent single vowels: one for the first syllable, the second for a »half-long» vowel following a (C)VC (short syllable) sequence which is longer than the first syllable vowel duration and the third for the V_2 position of other structures. In a more precise interpretation one might be justified in counting the modifying coefficients of the original duration for each vowel and consonant position in each of the eight bisyllabic basic structures presented in figure 7. The generation of three and more syllabic structures would obviously be possible by combining and modifying the bisyllabic patterns and their parts.

As far as the possibilities for automatic identification are concerned, on the other hand, the distribution of segmental duration of Finnish words forms a very complicated set of problems because of the multitude of factors affecting segment duration and also because the contrastive durational ratios important in identification change in a relatively complex way in different word structures under the influence of, for example, tempo and stress.

[1] Lindblom (1968) has described the duration of a vowel in Swedish by the mathematical formula $V = (d + D/n)k$ where d represents a constant duration characteristic of unstressed vowels in initial and medial positions within the word and D represents an increment due to superimposed stress and n the number of syllables. k is the position-dependent factor ($>$ 1 in final position: 1 otherwise) and D/n the factor of isochrony. The consequence of the last-named factor in Finnish was treated in chapter 3.9.

4.6. *The bandwidth of phonemic boundary*

The lines in figures 10_{1-29} describing the distribution of the listeners' answers, when the phonetic duration of different sound segments are changed do not descend steeply in any figure: a critical period appears in all the tests during which ever more listeners reassess the phonemic quantity of the changed segment or of some other sound in the stimulus, or in other words recognize the word heard as something else. Most of the curves are similar in structure: they start with a very gradual change and then continue very steeply and again end gently sloping. At a certain point the change in the duration of a sound segment causes more changes in word formation than around it. The time interval within which most of the changes in identification take place is called here the bandwidth of the phonemic duration boundary. The changing sound duration, which caused (or probably would have caused according to the graphic interpretation) as many single as double phoneme observations, is the theoretical centre of the phoneme boundary (cf Liberman 1961, p. 183; Abramson 1962, p. 109). For the measurement of the critical duration area or boundary area an arbitrary 'risk level' must be chosen. In this study 80 % of the changes were chosen as such. It is convenient, graphically, to distinguish the extreme 10 % bands from the answer lines and measure the time interval between the points of intersection of the 10 % horizontal lines and the answer lines. Measurement of the interval between the 100 % identifications would only have been possible in some cases and it would not have been possible to compare the results with corresponding tests performed in some other studies. Cases in the tests in this study where a change in a certain sound segment duration caused the reshaping of some other segment in the word structure have been omitted from the measurements of the bandwidth of the phonemic boundary as subject to various interpretations.

The critical duration bands of vowel distinctions in the various tests, i.e. the duration interval within which 80 % of the changes in the identification take place, are as follows:

single vowel/double vowel distinction in first syllable

tested word pair	80 % of changes within
etu/eetu	4.8 cs
uni/uuni	4.2 cs
esti/eesti	3.4 cs
udella/uudella	4.7 cs
	Mean 4.3 cs

single vowel/double vowel distinction in second syllable

tested word pair	80 % of changes within
taka/takaa	6.9 cs
mätä/mätää a)	4.2 cs
mätä/mätää b)	4.8 cs
pese/pesee	4.3 cs
kassa/kassaa	4.2 cs
takka/takkaa	7.6 cs
	Mean 5.3 cs

The cases where it has been possible to estimate the bandwidth of the phonemic boundary between durations of single and double consonants on the basis of the graphs are as follows:

tested word pair	80 % of changes within
taka/takka	7.5 cs
rysä/ryssä	10.0 cs
ripi/rippi	6.0 cs
takaa/takkaa	11.0 cs
mätää/mättää	7.4 cs
liikaa/liikkaa	9.9 cs
	Mean 8.7 cs

According to the means the critical band is slightly shorter in first than in second syllable vowels and noticeably narrower in distinction between vowel quantities than in distinction between single and double plosives and single and double fricative /s/. In the graphic descriptions of the tests one can also see that when the duration of consonants is changed the lines illustrating changes in observation may be more gently sloping than when the duration of vowels is changed. Since only phonetically long voiceless consonants /p t k s/ were used in the tests the results do not allow one to decide whether identification of consonant duration is in general more inaccurate than that of vowels, or whether this is merely the case with consonants which are long in intrinsic duration.

The identification tests of this study and the widths of boundary bands between single and double vowels measured on the basis of these tests can be compared to Abramson's (1962, p. 97, 105) corresponding tests administered with material from the Thai language. Abramson has used 25 or 50 msec stages as change intervals of vowel duration in his tests where the final double vowel members of the different minimal word pairs were to be changed. These steps seem to be too rough for an exact definition of the boundary and obscure the position of the boundary when it has fallen within a long transitional stage. However, the test results given by Abramson in the form of the tables are drawn here for comparison as graphs analogical to those used in the present study. In these the area of vowel duration within which 80 % perception changes took place is defined graphically in the same manner as above. The bandwidth between the average single and double vowel of thirteen of Abramsons's tests is 4.8 csec. The long stages in vowel duration between stimuli may, however, have affected the graphic definition of bands to some extent, but the values of Finnish and Thai are surprisingly similar, even though these languages are of completely different origin and of different absolute vowel durations.

There is no corresponding material for comparison of the perception boundaries of the duration difference between a single and double consonant. In the p/b test of Liberman et al. (1962, p. 183) 80 % of the changes in identification are situated approximately within a 50–55 msec interval and in Lisker's (1957, p. 46, fig 2) corresponding identification test approximately even within a 35 msec sequence. However, the durational difference between the categories of two consonant phonemes was under examination in these cases and not that between single and geminate consonants.

4.7. *Theory of the perception of phonemic length*

The tests which have been used in the present study do not provide a direct answer to the question of whether perception of duration at the phoneme boundary, that is, in the critical area between the durations of an accepted single sound and of an accepted double sound, is more precise than within the area of the phoneme. However, the preliminary perception tests seem to indicate that the absolute discrimination of durations is also clearly more precise at the phoneme boundary than inside a phoneme class: very small durational differences can be distinguished (even the 10 msec differences) in the critical area, but when those stimuli are under consideration where the duration of a sound varies within the accepted quantity degree, differences of corresponding size are not observed at all. The sharpened discrimination is also related to the fact that segments which are ambivalent in duration appear to be subjectively somewhat prominent (cf. Lehtonen 1969). Listeners' different abilities in discriminating between durational differences which are in some phonemic category, and those which are across the phonemic boundary, have been shown by Liberman & al. (1961 A). With the help of various tests made with synthetic speech, they have shown that the discrimination peaks of durations are situated in precisely that area where the change in duration (in their tests the change in the duration of the voiceless gap of the intervocal plosive) caused phonemic reinterpretation of a phonetic signal. This is the case, for example, with the synthetic words *rapid∼rabid* in their stimuli where the duration of the word medial plosive caused just as many /p/ as /b/ identifications, and which were, therefore, located on approximately the »phonemic» duration boundary. However, with the nonspeech stimuli, no appreciable increase in discrimination was found in the region corresponding to the location of the phoneme boundary. The discrimination peaks of speech stimuli were clearly higher and sharper than any peaks which appeared in these control data. It is slightly surprising, however, that no reduction of the discrimination was found within the phoneme category in the speech stimuli. The appearance of discrimination peaks on the phoneme boundary has also been indicated in the identification and discrimination tests of varying consonant transitions (Liberman & Harris & al. 1961 B). In these the researchers explain the appearance of discrimination peaks at the phoneme boundary of the speech stimuli on the basis of the so-called motor theory of speech perception: »We believe that in the course of his long experience with language, a speaker (and listener) learns to connect speech sounds with their appropriate articulations. In time, these articulatory movements and their sensory feedback (or more likely, the corresponding neurological processes) become part of the perception process, mediating between the acoustic stimulus and its ultimate perception» (Liberman & al. 1961 A, p. 177). In connection with transitions the appearance of discrimination peaks in the synthetic stimuli /b d g/ during listening were explained by referring to discontinuity of articulation:

b is a labial, *d* an apicodental and *g* a dorsovelar: »When significant acoustic cues that occupy different positions along a single continuum are produced by essentially discontinuous articulation (as, for example, in the case of second-formant transitions produced for /b/ by a movement of the lips and for /d/ by a movement of tongue), the preception becomes discontinuous and discrimination peaks develop at the phoneme boundary» (p. 177). This form of motor theory has been criticized by, for example, Malmberg (1969, p. 108–109). The former explanation cannot, however, be used, for example, as an explanation of the appearance of discrimination peaks in the p/b test, because there is no discontinuity between the articulatory movements of p and b. Nor can it be used in discrimination between vowel durations.[1]

According to the present author, how the listener is able to distinguish between non-linguistic differences in stimuli is not an essential point in speech perception, but the fact that the stimuli always tend to be categorized in some functional class. When the acoustic signal is ambivalent, for example, in those characteristics which affect identification between the expected /p/ and /b/ in a given context, the listener would automatically have to pay more attention to its acoustic characteristics (and possibly he would have to compare it to the various alternative articulations) in order to determine its phonemic function. If there is a tendency to show that the differences in the duration of plosive closure are more accurately observed when the durations are between the closure durations characteristic of /p/ and /b/ (c.f. Liberman & al. 1961 A, p. 177), it would be difficult to find any reason for the alleged absence of discrimination peaks between typical single and typical double vowel durations or between, for example, the Finnish single plosive /p/ and the geminate plosive /pp/. However, the existence of peaks in the discrimination of small physical differences between acoustic speech-signals is no condition of the

[1] Bastian and Abramson (1962, p. 743–744) have shown, by means of forced-choice discrimination tests, that Thai-speakers show no discrimination peak at the phoneme boundary. The test was made for the purpose of comparison with American subjects who behaved similarly to Thai-speakers. Fourteen duration variants were used for the test generated from the long member of a minimal pair of words by tape cutting. It is possible that the durational difference in the compared stimuli was, however, so great in the tests that the variation in the discrimination of absolute time intervals and in the discrimination of phonemically relevant durational differences did not appear. It is also possible, that the absolute, non-linguistic discrimination of durational differences in the periodic pulse signals (voiced sounds) is more accurate than the discrimination of the voiceless segments or nonperiodic noise-segments (plosives and voiceless fricatives) between voiced signals. For further discussion of identification and discrimination of phonic substance see also Lane 1966.

acceptance of the motor theory of speech perception.[1] In the special case of the perception of Finnish sound quantities the motor theory would mean that a listener does not immediately associate the received segment durations with single or double phonemes by comparing the auditory impression with some finite set of analogous reference patterns in the brain. Instead, it supposes that he has learned how the articulation mechanism as a whole and its various functions modify the timing of the signal intended to be produced. By means of this knowledge the listener is able to disclose the underlying motor timing-commands and to phonemize or decode the temporal patterns of the signal. The theory of speech perception with the transmission of a kind of motor servo-models, presented above, is, however, in close contact with the theories of speech production. Its validity will be unproved until we have more detailed information about these motor patterns of speech production which, according to the theory, would also function in the process of speech recognition.[2]

SUMMARY AND OUTLOOKS

The analyses of the present work were based on the speech of ten speakers with a homogeneous dialectal and social background. Thus, the analysis does not actually cover the whole area of standard Finnish but a geographically restricted subsystem – which in the opinion of the present author does not differ from the average representation in any essential respect. However, only a comparison of the differences in both production and perception of speech between speakers with different dialectal backgrounds will show the common relevant characteristics of the Finnish quantity patterns.

The analysis results as well as the results of the perception tests coincide in broad outline with the phonemization of long quantity as a sequence of

[1] The motor theory does not mean here any discontinuity of articulation, but the possibility that a listener maybe does not directly associate the acoustic qualities he receives with linguistic units like phonemes, syllables and words. Instead, according to the theory, he first interprets his auditory percepts in terms of the articulatory movements needed to produce these sounds and, in a second stage, recognizes the language units by association with these articulatory movements (Denes 1965, p. 252; See also Fay 1966, p. 99).

[2] Some aspects of the neuromuscular correlates of speech production are discussed by Lenneberg 1967, p. 89–120. The relations between the characteristics in the phonetic output of speech and the various innate mechanisms in the human vocal apparatus and auditory perceptual system are recently discussed also by Lieberman (1970).

two identical segmental phonemes. The sequence of phonetic segments is not, however, identifiable with the string of phonemes because of the various modifications of speech output caused by the production mechanism and used as identification cues in the perception. The factors affecting the durational patterns were grouped in paradigmatic (related to the intrinsic characteristics of various segmental phonemes) and syntagmatic (related to the higher-level units constituting of combinations of vowel and consonant segments). Amongst the former there were found, for instance, a compensation phenomenon between vowel and consonant segments: an intrinsically long consonant segment tends to shorten the surrounding phonetic sound segments and similarly, the consonant segment intrinsically short in duration has a lengthening effect on the preceding and following segments. In this way the total duration of the larger structure is kept rather constant independent of the qualities of the segmental sounds. Phenomena like this could be explained by a presumption of a certain hierarchy in the timing of phonemic stretches: the time axis is primarily phrased in terms of measures and in the second stage the time interval of a measure is filled with the segmental sounds of the actual word.

The perception tests with modified durations showed strikingly the existence of a unit larger than an individual sound segment as a reference frame of the perception of phonemic quantity. What a listener seems to identify when hearing a speech stimulus is not a queue of independent segments the phonemic quantities of which are recognized one by one in the order they reach his ear. The fact that the categorization of the quantity of a given sound segment is highly dependent on the changes in duration of the following segments is evidence of the function of syntagmatic patterns in speech perception. However, the tests performed with isolated bisyllabics enlighten just a part of the problem. An isolated word is just a simplified form of utterance for the purpose of testing. Actually the tests show that the evaluation of a given sound segment depends on the temporal relationships in the whole word, but they do not prove that a phonemic word itself would be the only relevant frame of its identification in an utterance containing several words. An aspect which has been left unnoticed in the present work is the relationship between temporal word patterns and internal juncture, which has been discussed by Lehiste (1965). There are cases in Finnish where the temporal structure of the word itself functions as a boundary signal, but, on the other hand, there may be many constructions where the distribution of duration in a sequence of phonetic sound segments does not give any information of word boundaries. Another question awaiting further investigation is that of stress and duration. The linguistic stress is partly manifested in lengthening of the stressed measure, but the significance of this and the other phonetic parametres (like pitch, timbre and possibly also intensity) can only be tested with perception tests based on synthetic material.

There are two characteristics in Finnish which justify one in calling it a 'quantity language'. Firstly, the name is justified because of the effective use of sequences of identical phonemes besides the combinations of dissimilar phonemes, secondly because of the important role the durational relationships have in the recognition of the various phoneme syntagms. However, though it has been shown that the changes in durational patterns alone can contribute to the identification of various quantity structures, it has not been proved that some other features (like the quality of a vowel) would not have any effect on quantity perception. On the contrary, it is possible that in cases of ambiguous stimuli or reduced signal, information for the identification of the word structures is also sought among the normally concomitant features of the full signal.

REFERENCES

ABERCROMBIE, D. (1967), *Elements of General Phonetics*. Edinburgh.

ABRAMSON, A. S. (1962), »The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments.» *Intern. Journ. of American Linguistics,* 28,2.

AIRILA, M. (1912), *Äännehistoriallinen tutkimus Tornion murteesta*. Suomi IV, 12, I. Helsinki.

ARISTE, P. (1941), »Hiiu murrete häälikute kvantiteedist.» *Acta et commentationes Universitatis Tartuensis* BXLIX.5. Tartu.

– (1953). *Eesti keele foneetika*. Tallinn.

BASTIAN, J. & ABRAMSON, A. S. (1962), »Identification and Discrimination of Phonemic Vowel Duration.» *Journal of the Acoust. Soc. of Am.* 34, 5, 743–744 (A.).

BELASCO, S. (1953), »The Influence of Force of Articulation of Consonants on Vowel Duration.» *Journal of the Acoust. Soc. of Am.* 25, 5, 1015–1016.

– (1958). »Variations in Vowel Duration: Phonemically Conditioned?» *Journal of the Acoust. Soc. of Am.* 30, 11, 1049–1050.

BENEDIKTSSON, H. (1963), »The Non-Uniqueness of Phonemic Solutions: Quantity and Stress in Icelandic.» *Phonetica* 10, 133–153.

BENNET, D. C. (1968), »Spectral Form and Duration as Cues in the Recognition of English and German Vowels.» *Language and Speech* 11,2, 65–85.

BOLLA, K. (1968) ВЛИЯНИЕ О КРЧЖЕНИЯ НА ДЛИТЕЛЬНОСТЬ ТЛАСНЫХ В РУССКОЙ РЕЧИНН *Annales Universitatis Scientiarum Budapestiensis de Rolando Eötvös nominatae.* Sectio Philologica, Tomus VIII, 35–50.

COLLINDER, B. (1928), »Über den Begriff der Quantität.» *Journal de la Société finno-ougrienne* XLII, 1–28.

– (1937), »Über Quantität und Intensität.» *Neophilologische Mitteilungen* 1937, 97–120.

– (1964), »Das Wort als phonetische Einheit.» *Sprachwissenschaft und Wahrscheinlichkeit.* Acta Universitatis Upsaliensis. Studis Uralica et Altaica Upsaliensia 1. Uppsala. (orig. published 1939).

DELATTRE, P. (1962), »Some Factors of Vowel Duration and Their Cross-Linguistic Validity.» *Journal of the Acoust. Soc of Am.* 34, 8, 1141–1143.

DELATTRE, P. & HOHENBERG, MARGARET (1968), »Duration as a Cue to the tense/lax Distinction in German Unstressed Vowels.» *Int. Review of Applied Linguistics* VI, 4, 367–390.

DELATTRE, P. & MONNOT, M. (1968), »The Role of Duration in the Identification of French Nasal Vowels.» *Int. Review of Applied Linguistics* VI, 3, 267–288.

DENES, P. (1955), »Effect of Duration on the Perception of Voicing.» *Journal of the Acoust. Soc. of. Am.* 27, 4, 761–764.

– (1965). »On the Motor Theory of Speech Perception.» *Proceedings of the Fifth Int. Congress of Phonetic Sciences*. Basel–New York. p. 252–258.

DIETH, E. (1950), *Vademekum der Phonetik.* Bern.

DONNER, K. (1912), *Salmin murteen kvantiteettisuhteista.* Suomi IV, 9, II. Helsinki.
DURAND, MARGUERITE, (1946). *Voyelles longues et voyelles brèves.* Collection linguistique publiée par la Société de linguistique de Paris XLIX. Paris.
ELERT, C. C. (1965), *Phonological Studies of Quantity in Swedish.* Uppsala 1965.
ENKVIST, N. E. (1962), »The Choice of Transcription in Foreign-Language Teaching.» *Proceedings of the Fourth Int. Congress of Phonetic Sciences.* The Hague. p. 586–589.
v ESSEN, O. (1962), »Trubetskoys 'fester' und 'loser' Anschluss in experimentalphonetischer Sicht.» *Proceedings of the Fourth Int. Congress of Phonetic Sciences.* The Hague. p. 590–597.
FANT, G. (1967), »Acoustis Aspects of Speech.» *Readings in Acoustic Phonetics*, ed. Lehiste. Cambridge, Massachusetts p. 93–107.
FAY, W. H. (1966), *Temporal Sequence in the Perception of Speech.* Janua Linguarum. Series Minor Nr XLV. The Hague.
FINTOFT, K. (1961), »The Duration of some Norwegian Speech Sounds.» *Phonetica* 1961, 7, 19–39.
FISCHER-JØRGENSEN, ELI (1955). »Om vokallængde i dansk rigsmål.» *Nordisk tidsskrift for tale og stemme*, 15, 1, 33–56.
– (1964). »Sound Duration and Place of Articulation.» *Zeitschrift für Phon.* 17, 2–4, 175-207.
FLANAGAN, J. L. (1965), *Speech Analysis Synthesis and Perception.* Würzburg.
FLIFLET, A. L. (1962), »Einige Beobachtungen über Anschluss und Silbe.» *Proceedings of the Fourth Int. Congress of Phonetic Sciences.* The Hague. p. 610–615
– (1963) »Syllable Type and Syllable Perception.» *Phonetica* 10,3–4, 187–193.
GÅRDING, EVA (1967), *Internal Juncture in Swedish.* Travaux de l'institut de phonétique de Lund VI. Lund.
HADDING-KOCH, KERSTIN & ABRAMSON A. S. (1964), »Duration versus Spectrum in Swedish Vowels: some Perceptual Experiments.» *Studia Linguistica* XVIII, II, 94–107.
HADDING-KOCH, KERSTIN & PETERSSON L. (1965), *Instrumentell Fonetik.* Lund.
HAKULINEN, L. (1961), *Suomen kielen rakenne ja kehitys.* Toinen, korjattu painos. Keuruu.
HARMS, R. T. (1960), »Stress and Juncture in Finnish», *American Studies in Uralic Linguistics.* Indiana University Publications. Bloomington.
– (1968) Review of »Consonant Quantity and Phonological Units in Estonian.» by Ilse Lehiste. *Language* 44,2, 407–409.
HEGEDÜS, L. (1959), »Beitrag zur Frage der Geminaten.» *Zeitschrift für Phon.* 12, 1–4, 68–106.
HENRYSSON, S. (1965), *Elementär statistik.* Uppsala.
HOCKETT, C. F. (1955), *A Manual of Phonology.* Baltimore.
HOUSE, A. S. (1961), »On Vowel Duration in English.» *Journal of the Acoust. Soc. of Am.* 33,9, 1174–1178.
HUGGINS, A. W. F. (1968), »The Perception of Timing in Natural Speech I.» *Language and Speech* 11, 1, 1–11.
ITKONEN, T. (1964), *Proto-Finnic Final Consonants I:1.* Helsinki.
– (1966). Nykysuomen kahden kasvot.» *Nykysuomen tutkimusta ja huoltoa* (SKS Tietolipas 48) Helsinki.
– (1968). »Itämerensuomalaisen liudennuksen fonologinen paradoksi.» Fenno-Ugrica. *Juhlakirja Lauri Postin 60-vuotispäiväksi.* SUST 145. Helsinki.
JAKOBSON, R. & FANT, G & HALLE, M. (1955), *Preliminaries to Speech Analysis.* Massachusetts Institute of Technology.
JONES, D. (1950), *The Phoneme: Its Nature and Use.* Cambridge.
KAISER, L. (1964), »Phonetic Similarity apart from Linguistic Affinity.» *Zeitschrift für Phon.* 17,2–4, 243–249.

KARLSSON, F. (1969), »Suomen yleiskielen segmentaalifoneemien paradigma.» *Virittäjä* 1969, 4, 351–362.
– (1970), *Finskans diftonger och vokalkombinationer*. An unpublished licentiate thesis in the Departement of Phonetics, University of Turku.
KIERIMO, U. (1929), »Äänteitten kestosuhteista.» *Suomen aistivialliskoulujen lehti* 38,5, 48–51.
LANE, H. (1966), »Identification, Discrimination, Translation: the Effects of Mapping Ranges of Physical Continua onto Phoneme and Semene Categories.» *Intern. Rew. of Appl. Linguistics* IV, 4, 215–226.
LAUROSELA, J. (1922), *Foneettinen tutkimus Etelä-Pohjanmaan murteesta*. Helsinki.
LAZICZIUS, G. (1966), »Zur Lautquantität.» *Selected writings of Guyla Laziczius*. Ed. Thomas A. Sebeok. The Hague – Paris.
LEHISTE, ILSE (1960), »Segmental and Syllabic Quantity in Estonian.» *American Studies in Uralic Linguistics*. p. 21–82.
– (1965 A), »The Function of Quantity in Finnish and Estonian.» *Language* 41,3, 447–456.
– (1965 B), »Juncture.» *Proceedings of the Fifth Int. Congress of Phonetic Sciences*. Basel–New York. p. 172–200.
– (1966), *Consonant Quantity and Phonological Units in Estonian*. Indiana University Publications, Uralic and Altaic Series, vol 65. Bloomington.
– (1968), »Vowel Quantity in Word and Utterance in Estonian.» Congressus secundus internationalis fenno-ugristarum Helsinki 1965. Pars I. Helsinki. p. 293–303.
– (1970), »Temporal Organization of Spoken Language.» *CISRC Working Papers in Linguistics* No. 4. The Ohio State University. Columbus, Ohio. p. 95–114.
LEHTONEN, J. (1968), »*Konnotatiivi intonaatio keskisuomalaisessa yleiskielessä*.» An unpublished licentiate thesis in the Departement of Phonetics, University of Helsinki.
– (1969) »Huomioita kvantiteettien foneemirajoista ja subjektiivisista kestohavainnoista.» *Virittäjä* 1969, 4, 363–370.
LENNEBERG (1967), *Biological Foundations of Language*. New York–London–Sydney.
LIBERMAN & al. (1961 A) = LIBERMAN, A. M. & HARRIS, K. S. & EIMAS, P. & LISKER, L. & BASTIAN, J. »An Effect of Learning on Speech Perception: the Discrimination of Durations of Silence with and without Phonemic Significance.» Language and Speech 4, 175–195.
LIBERMAN & al. (1961 B) = LIBERMAN, A. N. & HARRIS, K. S. & KINNEY, J. A. LANE, H. »The Discrimination of Relative Onset-time of the Components of Certain Speech and Nonspeech Patterns.» Journal of Experimental Psychology 61,5, 379–388.
LIEBERMAN, P. (1967), *Intonation, Perception and Language*. MIT Press, Cambridge, Mass.
– (1970). »Towards a Unified Phonetic Theory.» *Linguistic Inquiry* 1. 3., 307–322.
LIIV, G. (1961), »Eesti keele kolme vältusastme vokaalide kestus ja meloodiatüübid.» Keel ja Kirjandus 1967, 7. 412–424: 1961,8, 480–490.
– (1962). »On the Acoustic Composition of Estonian Vowels of Three Degrees of Length.» Eesti NSV Teaduste Akadeemia Toimetised. XI köide. Ühiskonnateaduste seeria, Nr 3 p. 271–286.
LINDBLOM, B. (1967) »Vowel Duration and a Model of Lip Mandible Coordination.» STL–QPSR–4/1967 (KTH) Stockholm. p. 1–29.
– (1968) »Temporal Organization of Syllable Production.» STL–QPSR 2–3/1968 (KTH) Stockholm. p 1–5.
LISKER, L. (1957), »Closure Duration and the Intervocalic Voiced-voiceless Distinction in English.» *Language* 33, 42–49.
MAACK, A. (1949). »Die spezifische Lautdauer deutscher Sonanten.» *Zeitschrift für Phon.* 3,3–4, 190–232.

– (1951). »Die Variation der Lautdaued Deutscher Sonanten.» *Zeitschrift für Phon.* 5. 287–340.
MAGDICS, KLARA (1961), »Kürzung der unbetonten Vokale in der ungarischen Umgangssprache.» *Zeitschrift für Phon.* 14,1, 21–44.
MALMBERG, B. (1944), *Die Quantität als phonetisch-phonologischer Bergiff.* Lunds universitetets Årsskrift N. F. Avd. 1. Bd 41. Nr 2. Lund.
– (1949), Review of »Voyelles longues et voyelles brèves.» by Marguerite Durand. *Studia linguistica* III,1, 39–59.
– (1963), *Structural Linguistics and Human Communication.* Berlin–Göttingen–Heidelberg.
– (1967), *Lärobok i fonetik.* Lund.
– (1969), *Introduktion till Fonetiken som vetenskap.* Lund.
MC NEMAR, Q. (1963). *Psychological Statistics (third edition).* New York and London.
MEYER, E. A. (1903), *Englische Lautdauer.* Uppsala.
MEYER, E. A. & GOMBOCZ, Z. (1909), »Zur Phonetik der ungarischen Sprache.» *Le Monde Oriental* 2 (1907–08; publ. 1909), 2–3, 122–187.
PALOMAA, J. K. (1946), *Suomen kielen äännekestoista puhumaan oppineen kuuromykän ja kuulevan henkilön ääntämisessä.* Publicationes Instituti Phonetici Universitatis Helsingforsiensis N:o 8. Turku.
PENTTILÄ, A. (1963), *Suomen kielioppi.* Toinen, tarkistettu painos. Porvoo.
PETERSON, G. E. & LEHISTE, ILSE (1960). »Duration of Syllable Nuclei in English.» *Journal of the Acoust. Soc. of Am.* 32,6, 693–703.
PIKE, K. L. (1947), *Phonemics.* (11. printing 1968). Ann Arbor
PILCH, H. (1968), *Phonemtheorie. I. Teil.* Zweite, verbesserte Auflage. Basel–New York.
POHLINK, M. (1968), »Ein Messplatz für Sprachuntersuchungen, bestehend aus Zeitfilter und Sprachschalter.» *Zeitschrift für Phon.* 1968, 3–4, 335–344.
POSTAL, P. M. (1968), *Aspects of Phonological Theory.* New York, Evanston, and London.
POSTI, L. (1968), Über das Quantitätssystem im Estnischen.» *Congr. sec. int. fenno-ugr. Helsinki* 1965. Helsinki. p. 408–418.
RAPOLA, M. (1965), *Suomen kirjakielen historia I.* SKS:n toimituksia, 197. Helsinki.
– (1966), Suomen kielen äännehistorian luennot. SKS:n toimituksia 283. Helsinki.
RAVILA, P. (1961), »Kvantiteetti distinktiivisenä tekijänä.» *Virittäjä* 1961,4, 345–350.
SADENIEMI, M. (1949), *Metriikkamme perusteet.* SKS:n toimituksia, 236. Helsinki.
SEBEOK, T. A. (1944), »Phonemes and Orthography in Finnish.» *Acta Linguistica* IV, 130–135.
SHARF, D. J. (1962), »Duration of post-stress intervocalic stops and preceding vowels.» *Language and Speech.* 5,1. 26–30.
SOVIJÄRVI, A. (1937), »Foneettisia havaintoja Tornionjoen ja Pohjanmaan murteiden švaavokaaleista.» *Suomi V, 19.* Helsinki.
– (1944), *Foneettis-äännehistoriallinen tutkimus Soikkolan inkeroismurteesta.* Suomi 103. Helsinki.
– (1946), »Huomioita puherytmiikasta.» *Virittäjä* 1946 p. 117–129.
– (1957), »The Finno-ugric languages.» *Manual of Phonetics.* ed. Kaiser. p. 312–324.
– (1958), »Alustavia mittaushavaintoja suomen yleiskielen sanapainosta.» *Virittäjä* 1958, 351–365.
– (1966), »Suomen yleiskielen foneemien luontaisista vastakohtapiirteistä.» *Virittäjä* 1966, 2. p. 189–196.
SOVIJÄRVI, A. & PELTOLA, R. (1965), *Suomalais-ugrilainen tarkekirjoitus.* 4. painos. Publicationes instituti phonetici universitatis Helsingiensis 9. Helsinki.
STRAIN, J. (1969). »A Preliminary Report on Vowel Duration in Persian.» *Intern. Rew. of Appl. Linguistics* VII,3, 201–205.

SWADESH, M. (1937), »The Phonemic Interpretation of Long Consonants.» *Language* 13, 1–10.

TRUBETZKOY, N. S. (1938), »Die phonologischen Grundlagen der sogenannten 'Quantität' in den verschiedenen Sprachen.» *Scritti in onore di Alfredo Trombetti.* Milano. p. 155–174.

– (1939), *Grundzüge der Phonologie.* 4. Auflage 1967. Göttingen.

VANVIK, A. (1966), *A Phonetic Analysis of the Dialect of Trondheim.* Oslo.

WIIK, K. (1965), *Finnish and English Vowels.* Annales Universitatis Turkuensis. Series B Tom. 94. Turku.

– (1969), »Onko suomen yleiskielessä oppositiota /i/–/j/ ja /u/–/v/?» *Juhlakirja Paavo Siron täyttäessä 60 vuotta.* Vammala. p. 223–237.

WIIK, K. & LEHISTE, ILSE (1968), »Vowel Quantity in Finnish Disyllabic Words.» *Congr. sec. int. fenno-ugr. Helsinki* 1965. Helsinki. 569–574.

ZIMMERMAN, S. A. & SAPON, S. M. (1958), »Note on Vowel Duration Seen Cross-Linguistically.» *Journal of the Acoust. Soc. of Am.* 30,2, 152–153.

ZWIRNER, E. (1933), »Quantität, Lautdauerschätzund und Lautkurvenmessung (Theorie und Material)». *Archives Néerlandaises phonétique expérimentale.* Tome VIII–IX. La Haye. p. 236–246.

ZWIRNER, E. & ZWIRNER, K. (1937), »Phonometrischer Beitrag zur Frage der neuhochdeutschen Quantität.» *Archiv für vergleichende Phonetik.* 1,2, 96–113.

ÄIMÄ, F. (1914). *Phonetik und Lautlehre des Inarilappischen.* Helsinki.

ÖHMAN, S. E. G. (1966), »Perception of Segments of VCCV Utterances.» *Journal of the Acoust. Soc. of Am.* 40,5, 979–988.