AN EXPERIMENTAL STUDY OF CONSONANT GEMINATION IN IRAQI COLLOQUIAL ARABIC

By

GHALIB BAQIR MOHAMMED GHALIB (M.Phil., Leeds)

Thesis Submitted in Accordance with the Requirements
for the Degree of Ph.D.

The University of Leeds
Department of Linguistics and Phonetics
July, 1984

This Thesis is Dedicated to the Memory
of my Late Father

CONTENTS

Part One : Linguistics of Gemination and Literature Survey

Appendices:

ABSTRACT

This thesis is an experimental study of the single/geminate contrasts in Iraqi Colloquial Arabic (I.C. Arabic). It falls into two major parts.

Part One is the theoretical part of this study. It comprises two Chapters. Chapter One explains the linguistic significance of gemination in Arabic and sheds some light on the historical background of the phonetic studies initiated by the old Arab grammarians. The syllable structure and the stress patterns both in Classical Arabic and in I.C. Arabic are dealt with. Recent theoretical and experimental studies on Arabic gemination are reviewed. Chapter Two presents a general survey of the literature. The distinctions between long, double and geminate consonants are clarified. Theoretical and experimental studies on gemination in languages other than Arabic are also reviewed.

Part Two is the experimental part of this study. It comprises five chapters. Chapter Three is an initial investigation that illustrates the acoustic characteristics of single and geminate consonants used in isolated words as well as in contextual utterances. Chapter Four is a synthetic investigation of geminate consonants carried out with the help of a microcomputer. It investigates the roles played by the segmental durations in the perception of single versus geminate consonants in I.C. Arabic. The theory of 'categorical perception' is then discussed in relation to our data. Chapter Five provides a palatographic evidence

of the distinctions between geminate consonants and their single counterparts. It also reviews earlier palatographic studies and discusses the main advantages of the technique of palatography, both static and dynamic, in phonetic research. Chapter Six represents the major experimental work in this study. It investigates the articulatory, aerodynamic and acoustic aspects of the geminate/non-geminate consonants in I.C. Arabic. Two experiments are included in this chapter; the pilot experiment and the main experiment. Because a considerable number of variables are involved in the two experiments, the results have been subjected to statistical treatment, including the statistical test of the 'analysis of variance'. Chapter Seven is the final chapter in this thesis. It begins with the conclusions one can derive from this study and ends with our suggestions for further investigations.

The major conclusions that can be drawn from this study are the following:

1. duration is the overriding factor in distinguishing a geminate from a single consonant both perceptually and productively;
2. there is little or no evidence for 'rearticulation' in the production of geminate consonants; and
3. no compensatory adjustment in vowel duration was observed in vowels preceding the two types of consonant in word-medial position.

ACKNOWLEDGEMENTS

As a postgraduate student in the Department of Linguistics and Phonetics I would like to thank all members of staff who have in some way or another contributed to my training and education. I appreciate their encouragement and friendliness during the preparation of this study.

My most sincere thanks go to my two supervisors Dr. Peter Roach (Head of Department) and Mrs. Celia Scully who have equally and skilfully shared the supervision of this work. I have greatly benefited from their excellent guidance and their invaluable comments and suggestions. I much appreciate their readiness at all times to discuss every problem I have encountered in this investigation.

I also register my gratitude to Mr. David Barber for sparing some time to discuss points I have raised in Chapter One. His comments and suggestions were precious. I also appreciate his readiness to offer help whenever help was needed.

I would like to extend my thanks to Mr. Eric Brearley (Chief Technician) who has done his utmost in preparing experimental set-ups and enabling me to carry out my experiments in the Phonetics Laboratory. I also appreciate his patience, help and friendliness while performing the experiments. My thanks are also extended to Mr. Colin Baxter (Technician) for making excellent studio recordings.

I am grateful to Dr. O.S. Pickering (Assistant Librarian) for his help in providing me with a private study at the Brotherton Library during the entire period of the preparation of this work. My gratefulness is also extended to the staff of the inter-library loans for their understanding and patience in helping me to obtain all necessary publications only available at limited academic institutes.

My thanks are also due to the Departmental secretaries for their sympathy, help and friendliness.

Thanks are also due to Miss Susan Nemes for her typing efforts and for bringing this work to its final form.

I extend my sincere thanks to my colleagues who acted as subjects in some experiments of this study and for whom phonetic research is an unknown world. They were always more than ready to offer help.

I also acknowledge the patience and the great efforts of my wife and children and their continuous encouragement towards the completion of this work.

Last but not least, I would like to express my deep gratitude to the University of Basrah for granting me with a three-year study-leave without which this study could not have been accomplished.

G.B.M. Ghalib

List of Phonetic Symbols

The following are the phonetic symbols used to facilitate the reading of the test words as well as the examples mentioned in this study.

1. The Consonants:

Symbol	Description
p	a voiceless bilabial plosive.
b	a voiced bilabial plosive.
t	a voiceless denti-alveolar plosive.
ṭ	a voiceless denti-alveolar emphatic plosive.
d	a voiced denti-alveolar plosive.
ḍ	a voiced denti-alveolar emphatic plosive.
k	a voiceless velar plosive.
g	a voiced velar plosive.
q	a voiceless uvular plosive.
?	a glottal stop.
f	a voiceless labio-dental fricative.
θ	a voiceless interdental fricative.
δ	a voiced interdental fricative.
δ̣	a voiced interdental emphatic fricative.
s	a voiceless denti-alveolar fricative.
ṣ	a voiceless denti-alveolar emphatic fricative.
z	a voiced denti-alveolar fricative.
ʃ	a voiceless palato-alveolar fricative.
x	a voiceless uvular fricative.
γ	a voiced uvular fricative.
ħ	a voiceless pharyngeal fricative.

ξ	a voiced pharyngeal fricative.
h	a glottal fricative.
tʃ	a voiceless palato-alveolar affricate.
dʒ	a voiced palato-alveolar affricate.
m	a voiced bilabial nasal.
n	a voiced denti-alveolar nasal.
r	a voiced alveolar flap.
l	a voiced alveolar lateral.
w	a voiced labio-velar approximant.
j	a voiced palatal approximant.

2. The Vowels:

Symbol	Description
i	a short half-close front with lip spreading vowel.
ii	a long close front with lip spreading vowel.
ee	a long half-close to half-open front with lip spreading vowel.
a	a short half-open unrounded vowel.
aa	a long open front unrounded vowel.
oo	a long half-close to half-open back rounded vowel.
u	a short half-close back rounded vowel.
uu	a long close back rounded vowel.

Abbreviations and Other Symbols

Ac	Area of constriction.
Ac min.	Minimum cross-sectional area of constriction.
Cf.	Compare.
cm^2	Centimeter squared.
cm/sec	Centimeter per second.
cm^3/sec	Centimeter cubic per second.
csec	Centisecond.
dB	Decibel.
DLs	Difference limens.
EMG	Electromyography.
EPG	Electropalatography.
et al.	And the others.
f.	Feminine.
FO	Fundamental frequency.
F1	First formant.
F2	Second formant.
F3	Third formant.
H.P.	High pass.
Hz	Hertz.
ibid.	In the same reference.
I.C. Arabic	Iraqi Colloquial Arabic.
imp.	Imperative.
JND	Just-noticeable difference.
k	Constant.
kHz	KiloHertz.
loc.cit.	In the page mentioned.

L.P.	Low pass.
m.	Masculine.
msec	Millisecond.
n.d.	No date specified.
op.cit.	In the reference mentioned.
p	Probability.
Po	Intraoral air pressure.
ΔP	Pressure drop across the tongue constriction.
S.D.	Standard deviation.
T	T value of the t-test.
U	U value of the Mann-Whitney U-test.
Uo	Volume flow rate of air through the mouth.
"	Inch.
"/sec	Inch per second.
>	More than.
<	Less than.
≤	Equal to or less than.
*	Statistically significant (where p value is at > 0.01 and ≤ 0.05).
**	Statistically highly significant (where p value is at ≤ 0.01).
$\sqrt{\ }$	Square root.
/ /	Phonemic transcription.
[ ]	Phonetic transcription.

Abbreviations of journals and periodicals

Am.J.Psych.	American Journal of Psychology.
ARIPUC	Annual Report of the Institute of Phonetics, University of Copenhagen.
IRAL	International Review of Applied Linguistics and Language Teaching.
J.Acoust.Soc.Am.	Journal of the Acoustical Society of America.
J.Exper.Psych.	Journal of Experimental Psychology.
JIPA	Journal of the International Phonetic Association.
J.Phon.	Journal of Phonetics.
J.Phon.Soc.Jap.	Journal of the Phonetic Society of Japan.
J.Speech Hear.Dis.	Journal of Speech and Hearing Disorders.
J.Speech Hear.Res.	Journal of Speech and Hearing Research.

INTRODUCTION

Gemination is a significant feature not only in Arabic but also in many other languages. Long ago the old Arab grammarians realized the significance of gemination in their language and they subjected it to meticulous description and analysis. Until quite recently most studies on gemination in Arabic have been in the main purely phonological and, to the best of our knowledge, there is as yet no adequate phonetic investigation of this phenomenon in any Arabic dialect. It is the aim of this study to investigate experimentally the phenomenon of gemination in Arabic and describe its relevant phonetic features.

Within the framework of the present study we have exploited the use of a variety of instrumental techniques to investigate contrasts between single and geminate consonants in I.C. Arabic. The acoustic data were obtained by making mingograph traces from intensity-meter outputs and also by making spectrograms. The aerodynamic data were obtained by making simultaneous mingograph traces from intraoral pressure and volume airflow rate outputs; whereas the articulatory data were obtained by making mingograph traces from the inferred oral area outputs and also by analyzing palatograms of the studied items. The perceptual data were obtained by conducting experiments using synthetic speech.

The material of this investigation represents words widely used in everyday speech by educated Iraqi Arabic speakers. Although the author of the present study comes from the city of Basrah, the dialect he uses and that of the other subjects does not belong to any specific area in the country and it can be heard everywhere. It is, in actual fact, readily understandable by all native speakers of Iraqi Arabic. This is most likely due to the influence of the different media of communication, such as television and broadcasting stations. Despite the divergencies among the dialects now existing in Iraq, the style used in this study is the most familiar in all parts of the country since it is not strictly marked by outstanding regional characteristics. It is not surprising if some Iraqi Arabic speakers disagree with the pronunciation given to this or that word for it is a common fact that no two people pronounce exactly alike; and it is well known that most people are unreasonably critical of a pronunciation which they do not happen to use themselves. (MacCarthy, 1950).

Albeit this study is primarily based on experimental work, we have found it very useful to start with a phonological background of the phenomenon of gemination in Arabic so as to help the reader understand its phonological significance before discussing its phonetic features. We greatly benefited from the use of various statistical tests to estimate how significant our results are at every stage of this work. This helped us to evaluate the significance of our results and assess the reliability of our findings.

PART ONE

LINGUISTICS OF GEMINATION AND LITERATURE SURVEY

CHAPTER ONE

THE LINGUISTICS OF GEMINATION IN ARABIC

1.1 Historical Background of Arabic Phonetic Studies

The science of Arabic phonetics is as old as Arabic grammar, and Arabic phonetic studies started simultaneously with Korānic studies. The old Arab grammarians were the first among their people to realize the importance of investigating and studying the various co-existent sounds and sound features in their language. We find, for instance, al-Khalīl,[1] the celebrated grammarian, was the first to write a phonetically based dictionary in Arabic known as 'al-Ain'. He innovated his own phonetic system on which he organized the Arabic consonants in accord with their place of articulation, beginning with consonants originated at the pharynx and ending with those produced at the lips. Sībawaihī (753-793) followed his teacher al-Khalīl in studying the sounds of Arabic. He was the first to present a comprehensive and systematic grammatical study of the Arabic language. In his famous book known as 'al-Kitāb', he presented a sophisticated classificatory organization of

---

[1] "Khalīl Ibn Aḥmad [Abū ᶜAbdurraḥmān Ul-Khalīl Ibn Aḥmad Ibn ᶜAmr Ibn Tamīm] (718-791), Arabian philologist, was a native of Oman. He was distinguished for having written the first Arabic dictionary and for having first classified Arabic metres and laid down their rules. He was also a poet, and lived the ascetic life of a poor student. His grammatical work was carried on by his pupil Sībawaihī. The dictionary known as the Kitāb-ul-ᶜAin is ascribed, at least in its inception, to Khalīl. It was probably finished by one of his pupils and was not known in Baghdad until 862. The words were not arranged in alphabetical order but according to physiological principles, beginning with ᶜAin and ending with Ya. The work seems to have been in existence as late as the 14th century, but is now only known from extracts in manuscript." (The Encyclopaedia Britannica, 11th ed., Vol. XV. Cambridge : CUP, 1911, p.770, 2nd column).

the articulation of sounds and some invaluable observations as to their description, assimilation and gemination. Sībawaihī analyzed the Arabic sounds as meticulously as he analyzed the language and for each sound or a group of sounds he indicated a particular place of articulation and gave it a precise description[1]. It is evident, then, that both al-Khalīl and Sībawaihī and their successors of the Arab grammarians classified and grouped the consonants in a way which tells us how far they recognized the contrasts emphasized by modern phonetics, such as between voiced and voiceless sounds, fricatives and plosives, etc. They realized that the formation of any sound essentially depends on either a close or an open approximation between two speech organs. Gairdner (1935) mentions this fact and points out that the Arab grammarians "notably anticipated modern phonetics in their classification of consonants as dental, palatal, velar, etc., and made the most exact observations as to the precise position of the tongue, palate, etc., associated with the production of the several sounds." (Al-Ani, 1978, p.187).

However, one can claim that these phonetic studies were purely descriptive; they resembled, to a great extent, the phonetic studies initiated by some grammarians in Europe during the seventeenth century. In this respect Cantineau (1960a)[2] states:

---

[1] For further details see Ghalib, 1977, p.47.

[2] Also see Cantineau, 1960b, pp.11-12.

"Les anciens grammairiens arabes ont été les premiers phonéticiens de leur langue : on trouve chez Sîbawaihi, par example, un classement correct des consonnes suivant leurs points d'articulation, des remarques importantes sur leurs modes d'articulation, une abondante étude de l' assimilation consonantique, des notions exactes sur la durée vocaligue et les alterations du timbre des voyelles, des indications sur les particularités phonétiques des différents dialectes. Comme celle de nos grammairiens du XVll^{e} siècle, cette phonétique des grammairiens arabes est purement descriptive, et ignore l'évolution historique de la langue; elle se borne à déclarer certaines prononciations correctes et d'autres vicieuses, sans aller au fond des choses. Elle n'en est pas moins fort précieuse, et bien des erreurs seraient évitees si l'on s'y reportait plu souvent." [p.5].

It goes without saying that the Arab grammarians paid much attention to studying the segmental entities in their language, particularly consonants which were subjected to extensive investigation. On the other side, less, or even no, attention was given to investigating the suprasegmental features, such as stress and intonation, until quite recently. Their failure in discussing such linguistic phenomena was not because they were unaware of their existence but, to our own knowledge, it was probably due to the fact that they regarded the sound segments as the vital carriers of intelligibility in oral communication and that the suprasegmental features were of marginal significance. Linguistic phenomena such as assimilation 'mumaaθala', elision 'haδf', emphasis 'tafxiim' and gemination 'taʃdiid' were amply studied as part of their grammar. At the same time, much ink was split in explaining the distinctions between voiced 'madʒhuura' and voiceless 'mahmuusa' sounds, and consonants were divided into two main groups in relation to these distinctions. The 'madʒhuura' consonants were described as "voiced, lenis, pressed, non-breathed and sonorous"; while their 'mahmuusa' counterparts as "voiceless, fortis, non-pressed, breathed and muffled." (Al-Ani, op. cit., p.103).[1]

Looking more narrowly at these descriptions one can note that the attention of the Arab grammarians was not mainly fixed upon the role of voice, but upon that of breath, while producing the 'madʒhuura' sounds. As reviewed by Gairdner (op. cit.), Howell (1883) closely followed the Arab grammarians to define the 'madʒhuura' consonants as "those letters[2] in which the current of breath is confined (munħabis, imprisoned) because, being strong in themselves and strong in the stress laid upon them in their outlet (maxradʒ), they are uttered only with a strong, hard sound, and the breath is pevented from running on with them." (Al-Ani, op. cit., p.188). It is quite obvious that the word 'voice' is not mentioned in this definition and much stress is laid on the action of the 'breath'. The 'mahmuusa' consonants were defined as "all unchangeably unvoiced sounds." (Al-Ani, op. cit., p.189).

---

1. More information on the distinctions between these terms is available in the well-researched article by Blanc, 1967.

2. The old Arab grammarians used to use the word "letter", i.e. 'ħarf', to mean consonant sound.

They were 'unvoiced' in terms that they were accompanied by more breath and lack of sonority.

1.2 Syllable Structure and Stress Patterns in Arabic

1.2.1 General Background

Before examining the syllable structure and the stress patterns in I.C. Arabic, we find it, as many other scholars do, necessary to refer first to the syllable structure and the stress patterns in Classical Arabic, though we have to bear in mind that a general statement about the structure of the syllable and the position of stress in a given Arabic dialect "should be based on an examination of the facts of that dialect, not on an assumed rule for an earlier stage of the language." (Ferguson, 1957, p.475). It is, in our opinion, important that a synchronic study of any Arabic dialect should be mainly based on the available diachronic data of Classical Arabic, since the latter has always been considered as the standard norm to which all Arabic dialects refer,[1] and since these different dialects are but cognates of their parent, namely Classical Arabic.

A common feature of Classical Arabic is that the isolate word may be pronounced in two forms; either with pausal or non-pausal forms. The pausal form, as regards the non-pausal form, involves principally the omission of the final inflectionals 'a', 'i', 'u', 'un', 'an' and 'in', and the use of a common ending 'ah' to correspond variously to the non-pausal '-atun/-atan/-atin'.[2] For example:

---

[1] Cf. Hassan, 1981, p.10.

[2] See Mitchell, 1975, p.77, and p.94 (fn.8) for further information.

Pausal Form	Non-Pausal Form	Meaning
ξallam	ξallama	he taught
juξallim	juξallimu	he teaches
muξallim	muξallimun	a teacher (m.)
muξallimah	muξallimatun	a teacher (f.)

The structure of the syllable, or its pattern, is describable in terms of the phonemic segments that make it up. (Harrell, 1962; Beeston, 1970; Al-Ani, 1978). These segments occur in certain recurrent patterns particular to the phonological distribution system of a given language. (Al-Ani and May, 1973). The structure of the syllable in Arabic is, undoubtedly, based on the phonemic system of Arabic. It is a well-known fact that Arabic is an inflectional language. This is to say, the grammatical relations are expressed by means of inflection as well as word order. Most of the 'roots' in Arabic are triconsonantal, e.g. /f-ξ-l/. They bear no lexical meaning without being integrated with 'infixes' which are usually composed of one or more vowels, e.g. /-a-a-/, which, in turn, together with the root /f-ξ-l/ form the 'stem' /faξal/. The stem in Arabic is regarded as the essential linguistic unit that carries the lexical information. Changes of the vowel distribution within the stem normally leads to grammatical and semantic changes, e.g. /faξal/ 'he did', but /faaξil/ 'a doer', /fiξil/ 'deed', etc. Other grammatical and semantic alterations can be brought about by the addition of certain prefixes and suffixes to the stem. The stem may have one or more prefixes and suffixes, e.g. /sajafξalu/ 'he will do', /faξaltu/ 'I did', /ˀafξaaluhum/ 'their deeds', /sajafξaluuna/ 'they will do', etc.

It is worth mentioning here that despite being a cognate of Classical Arabic, I.C. Arabic makes no distinction between pausal and non-pausal forms of words. In I.C. Arabic words are always pronounced with a distinct pausal form whether they occur in isolation or within a phrasal context. Besides, the word structure in I.C. Arabic has undergone some modifications which affect the phonetic values of some consonants and vowels. These modifications may affect the stem as well as the prefixes and suffixes of a word. For example:

Classical Arabic		I.C. Arabic	Meaning
Pausal	Non-Pausal		
kalb	kalbun	tʃalib	a dog
faqr	faqrun	fugur[1]	poverty
ṣaγiir	ṣaγiirun	ziγiir	small
qult	qultu	gilit	I said
samiξ	samiξa	simaξ	he heard
daras	darasa	diras	he studied
jaquul	jaquulu	jguul	he says
?aξmaak	?aξmaaka	ξimaak	he made you blind

In so far as stress is concerned, the old Arab grammarians did not adequately explain stress in Classical Arabic nor did they mention any phenomena which could be interpreted as stress despite their being meticulous in studying their language thoroughly. It seems to be that their

---

[1] Also heard as /fugur/ which may signify the adj. 'poor' as well as the n. 'poverty'.

failure to deal with stress was due to their feeling that stress was non-contrastive in Arabic, and that a change in the position of a word-stress would not bring about any change in its meaning. Therefore, they did not state any specific rules for stress patterns in the Arabic words. Ferguson (1956) comments on this point and says that "Classical Arabic had no word stress at all, either phonologically significant or automatic as a function of the syllabic structure of the word; that a pattern of fixed stress developed in Eastern non-Arabian Arabic ... and spread to many other dialects; and finally that this pattern was applied to Classical Arabic so that the present systems of stress used in reading Classical Arabic are later and derivative, not original." [pp. 384-385].

Nevertheless, the present stress rules in Classical Arabic have been largely based on how words in the holy Korān are recited. These rules, which are very often formulated in grammar books, do not take account of the fact that Classical Arabic is spoken differently in different regions of the Arab world, and they also appear to "take no cognizance of regional differences of pronunciation and, what is worse, to bear little or no resemblance to the facts in any one region." (Mitchell, op. cit., p.76).

Some recent studies (e.g. Birkeland, 1954; Ferguson, 1957; Harrell, 1957; Nasr, 1960) have greatly exaggerated the importance of the role played by word-stress in current Arabic dialects. These studies assume that some modern Arabic dialects have tendencies towards a new system in which stress has phonemic significance "which may be due to secondary developments within the respective dialects." (Birkeland, op.cit., p.8). They consider stress as the prime dynamic factor in all changes that take place in the syllable structures. It is believed that word-stress in most current Arabic dialects which have been investigated is of minor importance and there is even no decisive evidence as to its significant location in individual words. (Ghalib, op. cit., p.52). According to Contineau (1960b), what we really feel is the presence of a sentence-stress rather than a word-stress. In discussing stress in modern Arabic dialects, he states "En réalité, dans la plupart des dialectes arabes, l'accent de mot est faible, et il n'est nullement prouvé que sa place dans le mot soit stable. On a plutôt l'impression d'un accent de phrase que d'un accent de mot." [p.120]. Ferguson (1956) adds that "Every consideration of Arabic word-stress which does not take phrasal stress patterns into account is bound to be incomplete and misleading." [p.387].

However, the study of syllable structures and stress patterns of Arabic, as stated in the following sections, should, by no means, be considered as a comprehensive one. The main objective behind this attempt is to show, whenever examples are possible, the significance of the role played by geminate consonants within the structure of the syllable and the effect geminates may have on the location of stress in an Arabic word, whether Classical or dialectal.

1.2.2 Syllable Structure and Stress Patterns in Classical Arabic

1.2.2.1 Syllable Structure in Classical Arabic

In general, two types of syllable have been distinguished in Classical Arabic. This distinction mainly depends on the final segmental elements that characterize each syllable. Those syllables terminating either with a short vowel element (V), or with a long vowel element (VV) are called open syllables, and those syllables terminating with one consonant element (C), or perhaps two (CC), are called closed syllables.[1] For instance, the verb /faƐala/ 'he did' comprises three open syllables occurring successively, i.e. /fa + Ɛa + la/; whereas the noun /fiƐlun/ 'deed' is made up of only two closed ones, i.e. /fiƐ + lun/. Classical Arabic has a tendency towards having more words of closed syllables than words of open syllables. (Anīs, 1971).

Phonologically speaking, a syllable in Classical Arabic is characterized by the fact that it never begins with a vowel, and that the vowel may be preceded by one, and only one, consonant[2] and followed by one or two consonants, whether they are identical or non-identical. The vowel is regarded as the dominant element that determines the number of syllables. Consequently, by counting the number of vowels one can easily and automatically derive the number of syllables in a Classical Arabic word.

---

[1] Consonant clusters comprising identical or non-identical consonants are also represented by (CC).

[2] Word-initial geminates are considered special cases.

In accordance with their length, syllables in Classical Arabic may be divided into three categories:

1. The first category includes short syllables consisting of (CV) elements, e.g.

/bi/	'by, with'
/li/	'to, for'
/wa/	'and'

2. The second category includes medium syllables consisting of either (CVV) or (CVC) elements, e.g.

/fii/	'in'
/maa/	'what'
/ξan/	'about'
/min/	'from'

3. The third category includes long syllables consisting of either (CVVC) or (CVCC) or (CVVCC) elements. The final (-CC) elements may either represent identical or non-identical consonants, e.g.

/ξiid/	'festival'
/baaz/	'falcon'
/darb/	'path, road'
/ʃarq/	'east'
/sadd/	'dam'
/ħaadd/	'sharp, acute'.

Syllables in the first and second categories are the most frequently used in Classical Arabic. They may occur word-initially, word-medially and word-finally. Those in the third category are less frequent and they usually occur at the end of words spoken with pausal endings. The word in Classical Arabic may contain any of the above syllable patterns. For instance, the word /juqiimuun/ 'they reside' consists of the three different syllable categories; the first syllable /ju-/ is of the first category, the second /-qii-/ of the second category, and the third /-muun/ of the third category.

Modern Arab scholars (e.g. Ḥassaan, 1955; Bishr, 1970; Anīs, 1971) suggest that the word in Classical Arabic, whatever suffixes or prefixes it may have, does not comprise more than seven syllables. It is true that the vast majority of the Arabic words do not consist of more than four syllables, but there exist in Classical Arabic words that comprise even more than seven syllables. They may, in fact, contain as many as ten syllables. This can be attested by examining the syllable sequences in Classical Arabic words. Therefore, in Classical Arabic we may have words of:

1.	one syllable	as in /dʒahl/	'ignorance'
2.	two syllables	as in /dʒaahil/	'ignorant'
3.	three syllables	as in /jadʒhaluun/	'they are not aware'
4.	four syllables	as in /jatadʒaahal/	'he disregards'
5.	five syllables	as in /jatadʒaahaluun/	'they disregard'
6.	six syllables	as in /sajatadʒaahaluun/	'they will disregard'
7.	seven syllables	as in /jatadʒaahaluunahum/	'they disregard them'
8.	eight syllables	as in /sajatadʒaahaluunahum/	'they will disregard them'

9. nine syllables as in /sajatadʒaahaluunahumaa/ 'they will disregard them'[1]

10. ten syllables as in /fasajatadʒaahaluunahumaa/ 'they will disregard them'[1]

Monosyllabic words may take the form of any of the three syllable categories mentioned earlier. They may either be short, medium or long. The majority of short - and medium - syllable words are functional, i.e. grammatical, as in the two prepositions /fii/ 'in' and /min/ 'from', and the conjunction /wa/ 'and'. On the other hand, monosyllabic lexical words very often come in the form of long syllables, as in /ħuzn/ 'grief', /badr/ 'full moon', and /baab/ 'door'. Words which fall under this category are usually characterized by being spoken with their pausal forms, i.e. with the absence of inflexional suffixes. Words of more than two syllables are generally formed by the addition of prefixes and/or suffixes, as in /sajafʕalu/ 'he will do', /sajafʕaluuna/ 'they will do', etc.

1.2.2.2 Stress Patterns in Classical Arabic

Generally speaking, there are two subclasses within word-stress in Classical Arabic. The first subclass is Primary stress which is always associated with a pitch change when the word is said in isolation. The second subclass is Secondary stress which is achieved, to a large extent, by stress alone - not normally associated with a change of pitch

---

[1] The object 'them' indicates duality in this context.

direction.[1]

Following previous works initiated by other scholars (e.g. Birkeland, 1954; Mitchell, 1975) the following rules can be set up to specify the location of the primary stress of the most prominent syllable of the isolate word in Classical Arabic:

1. Words of one syllable usually take a primary stress, e.g.

/'riiħ/	'wind'
/'ħuut/	'whale'
/'xawf/	'fear'
/'layl/	'night'
/'saamm/	'poisonous'

2. Words with ultimate long syllables take oxytonic primary stress, e.g.

/ʃa'diid/	'strong, stern'
/sik'kiir/	'drunkard'
/sidʒ'dʒaad/	'carpets'
/mas'ruur/	'delighted'
/sa'miξt/	'I heard'

3. Words with penultimate medium syllables take paroxytonic primary stress, e.g.

/'saaħir/	'wizard'
/'muuħiʃ/	'lonely, deserted'
/'wadʒhak/	'your face'
/'kallam/	'he talked to'
/'ʃaxxaṣ/	'he identified'

---

[1] Cf. Gimson, 1970, p.224.

4. Words whose last two syllables have the structures of either (CV + CVC) or (CV + CVV) take proparoxytonic stress, e.g.

/'saaʕadak/	'he helped you'
/'qaddamak/	'he introduced you'
/tu'ʕallimak/	'she teaches you'
/'ʃaahaduu/	'they saw'[1]
/'darrasuu/	'they taught'[2]

On the other hand, word secondary stress is usually found in words comprising more than two syllables. This type of word-stress is quite often used in a very deliberate style of speaking and normally it precedes the primary stress, e.g.

/ˌsaaʕad'naaka/	'we helped you'
/ˌjastafii'duun/	'they benefit'
/ musˌtaʃfa'jaat/	'hospitals'
/jataˌħaadʒa'dʒuun/	'they argue'
/tataˌqaaʕa'duun/	'you retire'

Finally, it is of interest to mention here that in Classical Arabic no word-stress can take place on whatever syllable before the pre-antepenultimate.

1.2.3 Syllable Structure and Stress Patterns in I.C. Arabic

The following analysis of the various syllable structures and stress patterns existing in I.C. Arabic is

---

[1] As compared with /ʃaaha'duuh/ 'they saw him'.

[2] As compared with /darra'suuh/ 'they taught him'.

primarily based on my own intuitive knowledge as a native speaker of this dialect and on everyday conversations I usually have with the members of my family and with my friends who are reading for postgraduate degrees at Leeds University. Some of the rules referred to in this study are extracted from previous studies that dealt with Iraqi Arabic (e.g. Erwin, 1963, 1969; Woodhead and Beene, 1967; Ingham, 1975; Ghalib, 1977).

1.2.3.1 Syllable Structure in I.C. Arabic

Previously (section 1.2.1), it has been mentioned that Classical Arabic is considered as the parent of all modern Arabic dialects. This signifies that phonological, morphological and syntactical features are shared between these dialects and their parent.

It is a fact that I.C. Arabic has suffered some linguistic modifications, especially in its phonological and morphological systems, under the influence of some foreign languages used in the neighbouring countries. However, these modifications do not significantly violate the basic linguistic principles that are still operative both in I.C. Arabic and Classical Arabic, but they, undoubtedly, give the former its specific linguistic flavour and make it outstandingly different from other current Arabic dialects.

In so far as the syllable structure is concerned, and its similarity to that of Classical Arabic, one can say that there are as many syllables in I.C. Arabic words as there are vowels, whether these vowels are long or short. Thus, there is one syllable in /ʃaaf/ 'he saw', two in /ʃaafni/ 'he saw me', three in /ʃawwafni/ 'he showed me', four in /ʃawwafhiyyaa/ 'he showed it to him', and five in /ʃawwafithiyyaa/ 'I showed it to him', etc. It is characteristic that the first syllable of a word in I.C. Arabic may either begin with one or two consonants. Syllables other than the first always begin with only one consonant. Accordingly, whenever there are two consonants in the middle of a word, whether the two consonants are identical or non-identical, the point of syllable division is always between the two. (Erwin, 1969, p.28).

The distinction between open syllables and closed syllables, cited in section 1.2.2.1, also applies to the structure of the syllable in I.C. Arabic. A syllable is rendered 'open' when it ends with a short or long vowel, and 'closed' when it ends with one or two consonants. In terms of their length, short, medium and long syllables are also distinguished in I.C. Arabic. Unlike Classical Arabic, the consonant element that initiates a syllable in word-initial position may either be one consonant (C) or two consonants (CC). The latter may either by a consonant cluster or a geminate consonant. The following are the most possible types of syllables in I.C. Arabic. In words of more than one syllable, the required syllable is underlined.

structure	example	meaning	description
1. CV	/baħar/	sea	short, open
2. CCV	/δbaħa/[1]	he slaughtered it	medium, open
3. CVV	/daafi/	warm	medium, open
4. CCVV	/zbaala/	rubbish	long, open
	/bbaabak/	in your door	
5. CVC	/fallaʃ/	he demolished	medium, closed
6. CVCC	/qird/[2]	monkey	long, closed
	/miħtarr/	feeling hot	
7. CVVC	/baabkum/	your door	long, closed
8. CVVCC	/ħaarr/	hot	long, closed
9. CCVVC	/ħṣaan/	horse	long, closed
10. CCVCC	/sfandʒ/	sponge	long, closed
	/ṣfarr/	he became pale	

In respect to frequency of their shapes, the first seven categories occur more frequently than the last three; the four open categories being the most frequent of all, and the closed ones being the least frequent. Syllables of the last category, namely (CCVCC), are rare. They are only found in very few words most of which are of non-Arabic origins. The (CC) elements that begin and/or end some syllable categories may represent identical or non-identical consonants. When they terminate a syllable, that syllable must occur word-finally.

As for their syllabic distribution, syllables in category 1 (CV), 3 (CVV), 5 (CVC) and 7 (CVVC) operate freely

---

[1] May also be heard as /δibaħa/.

[2] May also be heard as /qirid/.

in I.C. Arabic, i.e. they occur without restrictions, initially, medially and finally in a word. Those in category 6 (CVCC) and 8 (CVVCC) may either occur in word-final position or they stand alone in monosyllabic words. On the other hand, syllables in category 2 (CCV), 4 (CCVV), 9 (CCVVC) and 10 (CCVCC) always exist in word-initial position.

It is a characteristic of I.C. Arabic speakers to omit short vowels occurring in the first unstressed syllables of bisyllabic words having the structure (CV + CVVC). Words such as /smiin/ 'fat', /ḥbaal/ 'ropes' and /fnuun/ 'arts' are very commonly used for the Classical Arabic words /sa'miin/, /ḥi'baal/ and /fu'nuun/, respectively. Conversely, they tend to insert a short vowel between (CC) elements occurring finally in monosyllabic words having (CVCC) structures provided that the two elements are non-identical. Thus, a monosyllabic word in Classical Arabic having the structure (CVCC) becomes a bisyllabic word in I.C. Arabic of (CV + CVC) structure. For example, the Classical Arabic words /ʃams/ 'sun', /farq/ 'difference' and /darb/ 'path' are very frequently heard in I.C. Arabic as /'ʃamis/, /'faruq/[1] and /'darub/, respectively.

There is also a tendency among I.C. Arabic speakers to lessen the number of syllables in words consisting of three or more syllables in Classical Arabic. Very often this lessening in the number of syllables is accomplished by eliminating the vowel of the second syllable of the word. Other changes may also occur, such as changing the vowel quality

---

[1] May also be heard as /'fariq/.

of the first syllable and/or the last syllable. For example, the I.C. Arabic words /'simξat/ 'she heard', /'simξaw/ 'they heard' and /'simξuu/ 'you hear (imp.)' are used for /'samiξat/, /'samiξuu/ and /'ʔismaξuu/ in Classical Arabic, respectively.

1.2.3.2 Stress Patterns in I.C. Arabic

Generally speaking, two types of stress pattern are distinguished in I.C. Arabic. The first type is determined by the syllabic structure of the word and the other by certain grammatical conditions. For simplicity of presentation, the first type will be referred to as type (1) and the other as type (2).

1.2.3.2.1 The Stress Pattern of Words in Type (1)

The stress pattern of this type is characterized by the presence of one primary stress as well as an optional secondary one in longer polysyllabic words. The presence of the secondary stress is a feature that characterizes a more careful speech style; while in a more rapid speech only the primary stress is clearly perceived. The location of the primary stress in this type is essentially dependent on the structure of the syllables in final and prefinal positions. The following rules are constructed:

1. Words of one syllable, no matter whether this syllable is short, medium or long, open or closed, usually take a primary stress, e.g.

/'min/	'from'
/'ṣeef/	'summer'
/'θuub/	'dress, shirt'
/'ṣaff/	'class, line'
/'ʃtamm/	"smell (v.)' or 'he smelt'

2. Words having the structure (CVVC) as their final syllables, take oxytonic stress, e.g.

/ṣa'diiq/	'friend'
/satʃ'tʃiin/[1]	'knife'
/tʃaδ'δaab/	'liar'
/maṭ'buux/	'cooked'
/dʒwaa'riib/	'socks, stockings'

The exception to this rule are those words in which the last two syllables have the structures (CVC + CVVC) or (CVVC + CVVC) of which the second syllable is always /-teen/, indicating duality in Arabic, very often take a paroxytonic stress, e.g.

/'santeen/	'two years'
/'ṣuurteen/	'two pictures'
/tif'faaħteen/	'two apples'
/ṣaa'buunteen/	'two soaps'

3. Words having the structure (CVVC) as their final syllable, also take oxytonic stress provided that the (CC) elements should constitute a geminate consonant, e.g.

---

[1] Also heard as /sitʃ'tʃiin/.

/miṭ'raʃʃ/	'he has become deaf'
/mix'rass/	'he has become dumb'
/miş'laξξ/	'he has become bald'
/mix'ḍarr/	'it has become greenish'
/miṣ'farr/	'it has become yellowish'

4. Bisyllabic words having the structures (CV + CV) or (CV + CVC) take a paroxytonic stress, e.g.

/'ʃinu/	'what'
/'ħilu/	'sweet, beautiful'
/'tʃiδib/	'lies'
/'ξalam/	'flag'
/'ʃuɣul/	'work'

5. Bisyllabic words having the structure (CV) or (CVC) as their final syllables and the structure (CVV) or (CVC) as their prefinal syllables, usually take a paroxytonic stress, e.g.

/'ħaafi/	'bare-footed'
/'xaayif/	'scared, frightened'
/'dʒiddi/	'serious'
/'maξmal/	'factory'
/'bilbil/	'bird'

6. Trisyllabic words having the structure (CV + CV) or (CV + CVC) as their last two syllables, very often take a paroxytonic stress. In such words there is frequently a non-contrastive free variation between paroxytonic and pro-paroxytonic stress, e.g.

or	/si'?ala/ /'si?ala/	'he asked him'
or	/fi'ʃaxa/ /'fiʃaxa/	'he wounded his head'
or	/fi'taħa/ /'fitaħa/	'he opened it'
or	/jif'tihim/ /'jiftihim/	'he understands' or 'he is clever'
or	/saa'farit/ /'saafarit/	'I travelled'
or	/naξ'ξasit/ /'naξξasit/	'I felt sleepy'

7. Polysyllabic words comprising not less than three syllables may take a secondary stress in addition to the primary one. The secondary stress often precedes the primary stress and is normally recognizable in a more careful speech style. Such words are usually of particular grammatical structures, e.g.

/ħaṭ,ṭeetilhij'jaa/	'I put it for him'
/sam,maξtilhij'jaa/	'I let him hear it'
/dʒi,bitlakij'jaaha/	'I brought it for you'
/ʃaw,waftakij'jaahum/	'I showed them to you'
/dʒi,maξitlakij'jaahum/	'I collected them for you'

1.2.3.2.2 The Stress Pattern of Words in Type (2)

Unlike the words in type (1), the words falling under this type are characterized by having their primary stress on syllables nearest to the beginning of the word. They are also characterized by the presence of certain grammatical particles, such as the interrogative particle /ʃ-/[1] 'what' and the three particles of negation /laa/, /maa/ and /muu/ 'not, no'[2] which are mainly used with verbs to denote negative commands or exhortations. With demonstrative pronouns, /laa/ and /muu/ are used to expound negative statements. (Ghalib, op.cit., pp.70-71).

e.g.

/'simaξ/	'he heard'
/'ʃsimaξ/	'what did he hear?'
/'ħitʃat/	'she talked'
/'ʃħitʃat/	'what did she talk?'
/jaak'luun/	'they eat'
/'ʃjaakluun/	'what do they eat?'
/'taaxuδ/	'you take'
/'laa taaxuδ/	'do not take'
/'maa taaxuδ/	'you should not take'
/'muu taaxuδ/	'I advise not to take'
/'haaδi/	'this'
/'laa haaδi/	'not this' (I advise you not to take one)
/'muu haaδi/	'not this' (This is not the one I am thinking of)

---

[1] The complete form of /ʃ-/ is /ʃinu/. The latter is less commonly used.

[2] By some speakers /laa/ and /maa/ are used with shorter vowel length, viz. /la/ and /ma/, respectively.

In addition to the above facts, the following rules are constructed:

1. When a noun is preceded by the preposition /bi/ 'in', it may either retain the primary stress in its normal position or the stress may be placed on the preposition. The variation of stress position in these words is non-contrastive, e.g.

	/bis'suug/	
or		'in the market'
	/'bissuug/	
	/bittaa'riix/	
or		'in history'
	/'bittaariix/	
	/bissaj'jaara/	
or		'by car' or 'in the car'
	/'bissajjaara/	
	/bil'baṣra/	
or		'in Basrah'
	/'bilbaṣra/	

2. When a verb or a demonstrative pronoun is preceded by the interrogative particle /'ʃinu/ 'what', the verb or the pronoun as well as the particle will take primary stress, e.g.

/'jiktib/	'he writes'
/'ʃinu'jiktib/	'what does he write?'
/'gaalat/	'she said'
/'ʃinu'gaalat/	'what did she say?'
/'δiitʃa/	'that' (f.)
/'ʃinu'δiitʃa/	'what is that?'

/δaa'kuul/	'those'
/'ʃinu δaa'kuul/	'what are those?'

3. When the particle /da/, which denotes continuity of verb tense, precedes a verb, the primary stress normally retains its normal position, e.g.

/'jiqra/	'he reads'
/da'jiqra/	'he is reading'
/'taakul/	'she eats'
/da'taakul/	'she is eating'
/ˌjilξa'buun/	'they play'
/daˌjilξa'buun/	'they are playing'

4. Like words in type (1), polysyllabic words of no less than three syllables may as well take a secondary stress in words of type (2). The secondary stress very often follows the primary stress and is usually heard in careful speech style, e.g.

/ʃi'kaalak/	'he complained to you'
/'maa ʃiˌkaalak/	'he did not complain to you'
/tin'ṭiihum/	'you give them'
/'laa tinˌṭiihum/	'do not give them'
/baa'goona/	'they robbed us'
/'maa baaˌgoona/	'they did not rob us'
/tikti'binna/	'you write to us'
/'laa tiktiˌbinna/	'do not write to us'.

1.3 The Linguistic Significance of Consonant Gemination in Arabic

Arabic consonants, like vowels, can be lengthened. This lengthening is usually referred to by the term 'gemination', i.e. doubling the consonant without an intervening vowel. Any Arabic consonant may occur geminate, especially in word-medial position, and in Arabic writing when it is particularly desired to indicate that a consonant is geminated, it is written only once and a relatively small sign like 'ّ', called 'ʃadda' or 'taʃdiid',[1] is placed over the consonant. However, this sign is not always inserted in Arabic texts despite its occasional vital importance for identifying a pattern.

Geminate consonants in Arabic were, and still are, generally orthographically represented with one letter only. It was al-Khalīl who was credited with the introduction of this sign which was merely an unlooped ʃ-sound derived from the word 'taʃdiid' itself. The introduction of the new sign was of prime importance so as to avoid confusion with corresponding words having single consonants since the difference between single and geminate consonants in Arabic is distinctive. Compare:

/'kasara/	'he broke'	/'kassara/	'he smashed'
/'kataba/	'he wrote'	/'kattaba/	'he made to write'
/'ħamala/	'he carried'	/'ħammala/	'he loaded'

---

[1] Both words are nouns and they signify 'reinforcement' or 'strengthening'.

The retaining of this sign in Arabic writing indicates that this device was preferred to doubling the letter since writing "would have been offensive to Arabic writers in so far as it involves a basic alteration in the received spelling and orthography." (El-Saaran, 1951, p.158). This means that this sign for gemination was intentionally devised to be significant in the sense that it is an abbreviation of the name of the phonetic feature characteristic of the consonant letter receiving the sign. The abbreviation may suggest that the consonant receiving the sign should be pronounced geminated.

One of the most crucial questions which has been put under extensive discussion for a long time is whether geminate consonants are to be treated as one long or two short segments.[1] According to Arabic grammar, the term geminate is applied to those consonants having long duration and determined by certain phonetic contexts, or to consonant letters receiving the sign representing such a long duration. El-Saaran (ibid., p.162) states that the old Arab grammarians regarded every geminate consonant sound as being equivalent to two sounds in measure and in pronunciation; the first being 'still' (or at rest) 'saakin', and the second 'moving' (or in motion) 'mutaharrik'.[2] He adds that, from the phonetic point of view, it is preferable to interpret a geminate consonant in Arabic as being double, i.e. "repetition of

---

1. This question will be discussed in full in languages other than Arabic in the following Chapter.

2. Both Wright (1962) and El-Saaran (1962) use the word 'quiescent' in parallel with 'movent' to mean 'still' and 'moving', respectively. Wright states that "A letter which has no following vowel is called 'a quiescent letter', as opposed to 'a movent letter'." [p.13]

sounds", and not long, i.e. having "true or indivisible length". He seems to follow Daniel Jones's interpretation in that intervocalic geminates should be considered as double and not single "on the ground that it is usually possible in precise speech to separate them into two by a diminution of force in the middle, attaching the first part to the first syllable and the second part to the second syllable." (Jones, 1967, p.116)[1]. His conclusion is based on the fact that in the syllabic division of Arabic, any intervocalic geminate consonant should be separated into two parts: part one is related to the first syllable and part two to the second syllable.

Cantineau (1960a) considers consonant gemination in Arabic to be equivalent to two identical single consonants, one occurring immediately after the other. For him a geminate consonant is only found in a context where a cluster of two non-identical consonants would be allowed, i.e. between vowels. He states:

> "... la géminée équivaut à deux consonnes simples identiques se suivant immédiatement: ce qui le prouve c'est qu'une géminée n'apparaît que dans la position où un groupe de deux consonnes serait admis, autrement dit entre voyelles;..." (p.189)

---

[1] Jones (1967) states that ... "In practice it is convenient to regard all intervocalic long consonants as double, on the ground that it is usually possible in precise speech to separate them into two by a diminution of force in the middle, attaching the first part to the first syllable and the second part to the second syllable. That long consonants are divisible in this way is obvious in the case of compound words and words formed with prefixes or suffixes, as in the following examples. English 'bukkeis (book-case), 'pennaif (pen-knife),... So also in Arabic where cognate words often have a vowel between the two consonants, e.g. dikka:n (shop), the plural of which is daka:ki:n (shops), or the Syrian mille (religion) which has plural milal." [pp.116-117]

He also adds:

> "Il n'y a pas ici de corrélation de gémination des consonnes, les géminées n'apparaissant que là où seraient admis des groupes de deux consonnes et une limite de mot ou de morphème pouvant passer entre les deux éléments de la géminée." (p.214)

Nevertheless, if we accept the definition of a geminate consonant as expressed by the old Arab grammarians and in terms of its syllabic structure; that it is a 'still' followed by a 'moving' sound, viz. having a (CCV) syllabic pattern, then a geminate in word-initial position is an impossible pattern in Arabic. A geminate in Arabic can then only occur either word-medially or word-finally. The (V) element of the geminate (CCV) can be any of the three short vowels [i], [a], or [u], as in /hajjinun/ 'easy, manageable', /sallama/ 'he greeted', and /taħassubun/ 'expectation', respectively. It can also be any of the long vowels [ii], [aa], or [uu], as in /ṣiddiiqun/ 'honest', /rukkaabun/ 'passengers' and /mutaħaabbuuna/ 'loving each other', respectively. Similarly, a geminate consonant can either be preceded by any of the three short vowels, as in /sittun/ 'six', /saddun/ 'a dam', and /ħurrun/ 'free, released', or by some of their long counterparts, as in /qaaṣṣun/ 'story-writer'. In word-final position, gemination is only possible when the word takes a pausal form where the geminate (CCV) pattern is reduced to (CC). For example:

non-pausal	pausal	meaning
/birrun/	/birr/	kindness
/sammun/	/samm/	poison
/ħubbun/	/ħubb/	love

However, geminate consonants occurring word-finally are non-distinctive in Arabic because contrasts between single and geminate consonants in this position are non-significant. An Arab, for instance, does not distinguish between /damm/ and /dam/. For him they both stand for one and the same word, namely 'blood', and while the first pronunciation seems to him rather exaggerated and deliberate, the second is quite normal and more acceptable. Pronunciation of a final geminate consonant requires a conscious effort to make it perceptible, as was recognized by the Arab grammarians. Cowell (1964) states that "In word-final position, any geminate consonant may occur after an accented vowel. At the end of a phrase, however, geminate consonants do not actually contrast with single ones; pronouncing, or even writing, them geminate simply serves to show the position of the accent and their potential significant length before vowels." [pp.23-24]. Johnstone (1967) agrees with Cowell's statement and contends that a final geminate in Arabic is characterized by tenseness of articulation as compared with a non-geminate.

One of the most interesting features of consonant gemination in Arabic is the way a geminate behaves in derivative situations. Arabic derivations often have a vowel inserted between the components of a geminate consonant. This situation can be clearly observed when forming plurals, and while in the singular form the geminate is retained, in the plural form it is separated by inserting a vowel, very often [a], between its components. For example:

singular		plural	
/muddatun/	'period'	/mudadun/	'periods'
/ξuddatun/	'equipment'	/ξudadun/	'equipments'
/sikkiinun/	'knife'	/sakakiinu/	'knives'

In other instances the reverse situation may take place. This is to say, the plural form takes the geminate consonant and the singular form keeps its single counterpart. Besides, other alterations in the preceding and/or following vowels and the location of stress may be noticed between the word in its singular form and that in its plural form, as is evident in these examples:

singular		plural	
/'ṭaalibun/	'student'	/ṭul'laabun/	'students'
/'kaatibun/	'writer'	/kut'taabun/	'writers'
/'ħaarisun/	'guard'	/ħur'raasun/	'guards'

Furthermore, in some instances the geminate consonant is found both in the singular form as well as the plural form. For example:

singular		plural	
/muṣawwirun/	'photographer'	/muṣawwiruuna/	'photographers'
/muṣammimun/	'designer'	/muṣammimuuna/	'designers'
/maħaṭṭatun/	'station'	/maħaṭṭaatun/	'stations'
/muξallimatun/	'teacher'	/muξallimaatun/	'teachers' (f.)

1.4 Consonant Gemination as a Result of Assimilation

1.4.1 In Classical Arabic

The phenomenon of assimilation has been among the most intensively studied features in the phonetics of Arabic by the old Arab grammarians. Their main interest behind studying this phonetic phenomenon, as well as many others, was because of the extreme care one should pay in the production of Arabic sounds while accurately reciting verses from the holy Korān, keeping in mind that the holy Korān is not only recited by native Arabic muslim speakers but also by millions of muslims for whom Arabic is a foreign language. Therefore, the way Arabic spoken by Arabic expert reciters of the holy Korān was, and it still is, regarded as the most accurate oral picture of Classical Arabic; the non-regionally biased style of spoken Arabic.

In certain phonetic contexts, gemination has been considered as a natural consequence of the phenomenon of assimilation. And whenever the expert reciters of the holy Korān feel it is inconvenient to assimilate one particular sound to another, which either precedes or follows it, they deliberately lengthen it and make it outstandingly perceivable so as to avoid its assimilation into the other. (Anīs, 1971, p.157). However, when they feel that one consonant is preferably assimilated to another consonant, the assimilation is phonetically, or even graphically, executed by the elision of the pausal sign, called 'sukuun', from the assimilated consonant. This very often applies to the l-sound of the only definite article in Arabic /ʔal/, usually when the latter precedes one of the solar consonants which are, almost all, originated at the alveo-dental zone.[1]

The Arabic definite article is normally prefixed to every noun that it limits without affecting its accentuation. For instance, when it limits a noun beginning with a solar consonant, the l-sound of the article is assimilated to the subsequent consonant which is in turn geminated in pronunciation. This can be illustrated in the following examples:[2]

without definite article			with definite article			
	word	meaning	word			meaning
[t]	tamr	dates	ʔattamr	or	ttamr	the dates
[d]	darb	path, road	ʔaddarb	or	ddarb	the path, the road
[ṭ]	ṭayr	bird	ʔaṭṭayr	or	ṭṭayr	the bird
[ḍ]	ḍayf	guest	ʔaḍḍayf	or	ḍḍayf	the guest
[δ]	θawr	bull	ʔaθθawr	or	θθawr	the bull
[s]	δanb	fault	ʔaδδanb	or	δδanb	the fault
[ṣ]	sayf	sword	ʔassayf	or	ssayf	the sword
[z]	ṣayf	summer	ʔaṣṣayf	or	ṣṣayf	the summer
[z]	zawdʒ	pair, spouse	ʔazzawdʒ	or	zzawdʒ	the pair, the spouse
[δ̣]	δ̣ulm	tyranny	ʔaδδulm	or	δ̣δ̣ulm	the tyrrany

---

1. The Arabic consonants are, on phonetic grounds, divided into solar (or sun) consonants and lunar (or moon) consonants. The terminology is, in fact, based on the fact that the Arabic word for sun /ʃams/ begins with a solar consonant and that for moon /qamar/ begins with a lunar consonant. The solar consonants are: [t, d, ṭ, ḍ, θ, δ, s, ṣ, z, δ̣, ʃ, n, l, r]. These consonants might form what is described in terms of modern phonological theory as 'a natural class', i.e. a group of homorganic consonants. Using the phonetic features suggested by Chomsky and Halle (1968), they can be defined as + consonantal, + coronal and + anterior. Consonants which are not included within this general definition are then the lunar ones. They are also fourteen in number. However, in Arabic writing, it should be noted that the definite article has to be written whether it precedes a solar or a lunar consonant.

2. All examples are monosyllabic words pronounced with their pausal forms.

[ʃ]	ʃarq	east	ʔaʃʃarq	or	ʃʃarq	the east
[n]	nuur	light	ʔannuur	or	nnuur	the light
[l]	lawn	colour	ʔallawn	or	llawn	the colour
[r]	rumħ	spear	ʔarrumħ	or	rrumħ	the spear

On the other hand, when followed by a lunar consonant, the l-sound of the definite article should not be assimilated into that consonant. Both the l-sound of the article and the subsequent lunar consonant must be pronounced in full. For example:[1]

	Without definite article		with definite article	
	word	meaning	word	meaning
[b]	baħr	sea	ʔalbaħr	the sea
[k]	kalb	dog	ʔalkalb	the dog
[q]	qayd	chain	ʔalqayd	the chain
[dʒ]	dʒisr	bridge	ʔaldʒisr	the bridge
[ʔ]	ʔanf	nose	ʔalʔanf	the nose
[f]	faʔas	axe	ʔalfaʔs	the axe
[x]	xawf	fear	ʔalxawf	the fear
[ɣ]	ɣarb	west	ʔalɣarb	the west
[ξ]	ξabd	slave	ʔalξabd	the slave
[ħ]	ħilm	dream	ʔalħilm	the dream
[h]	hams	whisper	ʔalhams	the whisper
[m]	mawt	death	ʔalmawt	the death
[w]	waqt	time	ʔalwaqt	the time
[j]	jawm	day	ʔaljawm	the day

In point of fact, the phonetic realization of an Arabic utterance contains noticeably more geminate consonants than the Arabic writing may suggest. One example is the way the

---

[1] All examples are monosyllabic words pronounced with their pausal forms.

l-sound of the definite article assimilates to the solar consonant sounds mentioned above, and another example is the way in which many geminations occur in written texts as a result of phonetic assimilations between the final consonant of one word and the initial consonant of the following word, or even between the end of one grammatical unit and the beginning of another unit. These assimilations and their resultant geminates simply have the effect of making identical two consonants which are already more or less related. We find that assimilation exists at places where there is a close syntactic link between the two neighbouring words. Quite often, examples have been quoted from the holy Korān and from classical literary texts. For example:[1]

/qur rabbi/	for	/qul rabbi/	
/jaγfil lakum/	for	/jaγfir lakum/	
/ˀidmix xaadimak/	for	/ˀidmiγ xaadimak/	'Give your servant a blow on the head'
/ˀislaγ γanamak/	for	/ˀislax γanamak/	'Skin your sheep'
/mim man/	for	/min man/	'From whom'
/ˀal laa/	for	/ˀan laa/	'Lest, so that not'
/maj jaquul/	for	/man jaquul/	'Who says'
/wa dʒaaˀas sajjaaratun/	for	/wa dʒaaˀat sajjaaratun/	
/laqadʒ dʒaaˀakum/	for	/laqad dʒaaˀakum/	
/qas saˀalaha qawmun/	for	/qad saˀalaha qawmun/	
/qaʃ ʃaγafaha ħubban/	for	/qad ʃaγafaha ħubban/	
/waman juriθ θawaaba ddunjaa/	for	/waman jurid θawaaba l dunjaal/	
/waˀiz zajjana lahum ʃʃajṭaan/	for	/waˀiδ zajjana lahum l ʃajṭaan/	
/wa?in tuξdʒif faξadʒabun/	for	/waˀin tuξdʒib faξadʒabun/	

---

[1] Examples_without their English translation are quoted from the holy Korān.

It is quite obvious that all the above examples show regressive assimilation, i.e. the final consonant of the preceding word is assimilated to the first consonant of the following word. Progressive assimilation also exists in Arabic, but it occurs less frequently.[1]

1.4.2 In I.C. Arabic

Most of the examples which have been discussed to show consonant gemination existing as a result of assimilation in Classical Arabic do not apply to I.C. Arabic. This is essentially because in everyday Iraqi speech different phrases are used to bear the same semantic interpretations. The examples quoted earlier are only limited to Classical Arabic.

However, in I.C. Arabic word-medial gemination occurring as a consequence of the assimilation of a voiceless consonant by a following voiced or voiceless consonant is of common prevalence. [t] in the following contexts is a good case in point.

[d] replaces [t] before [d], as in:

/middaaxliin/	for	/mitdaaxliin/	'they are entangled'
/middaaξmiin/	for	/mitdaaξmiin/	'they are collided'

[ḍ] replaces [t] before [ḍ], as in:

/miḍḍaamniin/	for	/mitḍaamniin/	'they are unified'
/miḍḍaarbiin/	for	/mitḍaarbiin/	'they are involved in a quarrel'

---

[1] For full details and extra examples see Anīs, op.cit., especially Chapter seven, pp.179-207; Anṭāki, pp.166-187; Cantineau, op.cit., pp.35-43.

[dʒ] replaces [t] before [dʒ], as in:

/midʒdʒaahliin/	for	/mitdʒaahliin/	'they disregard', 'they are not paying attention to'
/midʒdʒaawbiin/	for	/mitdʒaawbiin/	'they are respondent'

[ṭ] replaces [t] before [ṭ], as in:

/miṭṭaabqiin/	for	/mitṭaabqiin/	'they are correspondent'
/miṭṭalligiin/	for	/mitṭalligiin/	'they are divorced'

[s] replaces [t] before [s], as in:

/missaamħiin/	for	/mitsaamħiin/	'they are forgiving'
/missaahliin/	for	/mitsaahliin/	'they are lenient'

[ṣ] replaces [t] before [ṣ], as in:

/miṣṣaalħiin/	for	/mitṣaalħiin/	'they are on good terms'
/miṣṣaaħbiin/	for	/mitṣaaħbiin/	'they become friends'

[z] replaces [t] before [z], as in:

/mizzawdʒiin/	for	/mitzawdʒiin/	'they are married'
/mizzaaʕliin/	for	/mitzaaʕliin/	'they are on bad terms' 'they are angry with each other'

[ʃ] replaces [t] before [ʃ], as in:

/miʃʃaawfiin/	for	/mitʃaawfiin/	'they have met before'

In addition to words characterized by word-initial geminates resulting from the assimilation of the l-sound of the definite article /ʔil/ and a subsequent solar consonant, in a way similar to that mentioned in section 1.4.1, a geminate occurring word-initially may exist in some verbs which are characterized by a prefixed 't' and a word-medial geminate. The prefix 't' assimilates to a following consonant if the latter is [d], [ṭ], [ḍ], [tʃ], [dʒ], [θ], [δ], [s], [ṣ], [z], [ʃ], as in the examples below:

/ddarrab/	for	/tdarrab/	'he was trained'
/ṭṭaʃʃar/	for	/tṭaʃʃar/	'it was scattered'
/ḍḍamman/	for	/tḍamman/	'it included'
/tʃtʃabbaʃ/	for	/ttʃabbaʃ/	'he was tamed', 'he was broken in'
/dʒdʒammaξ/	for	/tdʒammaξ/	'it gathered'
/θθaqqaf/	for	/tθaqqaf/	'he was educated'
/δδakkar/	for	/tδakkar/	'he remembered'
/ssadʒdʒal/	for	/tsadʒdʒal/	'it recorded'
/ṣṣaddar/	for	/tṣaddar/	'it was exported'
/zzawwadʒ/	for	/tzawaadʒ/	'he got married'
/ʃʃammas/	for	/tʃammas/	'he sunbathed'

Such geminates may also occur in verbs with word-medial single consonants provided that the vowel which follows the geminate should be long. For example:

/ddaajan/	for	/tdaajan/	'he borrowed'
/ṭṭaawal/	for	/tṭaawal/	'he was impertinent'
/ḍḍaaξaf/	for	/tḍaaξaf/	'it was doubled'
/tʃtʃaalab/	for	/ttʃaalab/	'it was clung'
/dʒdʒaahal/	for	/tdʒaahal/	'he disregarded'
/θθaawab/	for	/tθaawab/	'he gaped'
/δδaabaħ/	for	/tδaabaħ/	'he fought fiercely'
/ssaalam/	for	/tsaalam/	'he shook hands with'
/ṣṣaaraξ/	for	/tṣaaraξ/	'he wrestled with'
/zzaaξal/	for	/tzaaξal/	'he became angry with'
/ʃʃaaʔam/	for	/tʃaaʔam/	'he was pessimistic'

On the other hand, the [l] of the determiner /hal-/, 'this, these' is assimilated to a subsequent consonant, particularly when that consonant is originated at the dental or alveo-dental, or palato-alveolar regions. For example:

/hattiin/	for	/haltiin/	'these figs'
/haṭṭiin/	for	/halṭiin/	'this clay'
/haddaar/	for	/haldaar/	'this house'
/haḍḍeef/	for	/halḍeef/	'this guest'
/hatʃtʃalib/	for	/haltʃalib/	'this dog'
/hadʒdʒaar/	for	/haldʒaar/	'this neighbour'
/haθθoor/	for	/halθoor/	'this bull'
/haδδiib/	for	/halδiib/	'this wolf'
/hassuug/	for	/halsuug/	'this market'
/haṣṣuuṭ/	for	/halṣuuṭ/	'this voice'
/hazzoodʒ/	for	/halzoodʒ/	'this pair'[1]
/haʃʃaab/	for	/halʃaab/	'this young man'
/haδ̣δ̣aalim/	for	/halδ̣aalim/	'this tyrant'
/hannuur/	for	/halnuur/	'this light, this brightness'
/harradʒul/	for	/halradʒul/	'this man'

There are other types of assimilation involving some consonants that result in gemination in everyday Iraqi speech. For instance the [l] sound of the preposition /ˀil-/ 'to, for' becomes [n] when the former is followed by [n], as in /ˀinna/ instead of /ˀilna/ 'to us'. The consonant [l] also becomes [n] in other words, such as /baddanna/ for /baddalna/ 'we changed', and /maanna/ for /maalna/ 'ours'. The [n] of the very widely-used prefix /dan-/, usually used to indicate continuity of action, is assimilated to an

---

[1] This expression may also be used sarcastically to mean 'this nincompoop'.

ensuing consonant to form a geminate, as in these contexts:

/darruuħ/	for	/danruuħ/	'let's go, we are going'
/dammuut/	for	/danmuut/	'let's die, we are dying'

The assimilated forms may also occasionally be heard, but in everyday normal-speed conversation the assimilated forms are of common use. (See Erwin, 1969, p.74). Similarly, the consonant [n] is pronounced [r] when it is followed by [r] and pronounced [m] when it is followed by [m], as in these examples:

/jirraad/	for	/jinraad/	'it is needed'
/marruuħ/	for	/manruuħ/	'we don't go'
/jimmurid/	for	/jinmurid/	'it is crushed'
/jimmilix/	for	/jinmilix/	'it is torn'.

1.5 Geminate Proper

1.5.1 In Classical Arabic

Earlier, in section 1.3, it has been mentioned that all consonants in Arabic, whether they are solar or lunar, are capable of being geminated in word-medial position. This gemination can function to distinguish one word from another, and contrasts can be made between the geminated consonants and their corresponding single cognates. It is this type of gemination that we usually refer to when we talk about geminate consonants in Arabic, and a geminate consonant of this category may be appropriately termed as 'geminate proper'.

Unlike gemination which occurs as a result of the process of assimilation, geminate proper is always a basic part of the internal structure of the word. It is an important factor in determining the syllabic pattern of words. (see section 1.2.2.1).

In Arabic, differences between words having medial geminate consonants and others having their corresponding single consonants are extremely significant, and there are pairs of words which are differentiated on the grounds that one contains a medial geminate consonant and the other contains its single counterpart. But this, of course, does not necessarily imply that every word in Arabic with a geminate consonant, whether it is word-initial or word-medial, has a corresponding word with a single consonant, or the other way round. The following are examples having intervocalic single consonants, both solar and lunar, with corresponding words containing their geminate cognates. They are presented in matching pairs to make comparison easier and more obvious. The chosen words are trisyllabic verbs having the patterns of (faξala) for the singles and (faξξala) for the geminates.

single		geminate	
/rasaba/	'he failed'	/rassaba/	'he caused to fail'
/naṣaba/	'he constructed'	/naṣṣaba/	'he appointed'
/ʃaɣala/	'he occupied'	/ʃaɣɣala/	'he offered a job'
/raħala/	'he departed'	/raħħala/	'he caused to depart'
/ξabara/	'he crossed'	/ξabbara/	'he made to cross'
/kataba/	'he wrote'	/kattaba/	'he made to write'
/raʔasa/	'he headed'	/raʔʔasa/	'he made to be the head'
/lamaħa/	'he saw'	/lammaħa/	'he indicated'
/ħalaqa/	'he shaved'	/ħallaqa/	'he flew'
/darasa/	'he studied'	/darrasa/	'he taught'

This contrast between geminate and non-geminate consonant does not only occur in verbs. It is also found in other parts of speech, such as nouns and adjectives. For example:

/dʒamaalun/	'beauty'	/dʒammaalun/	'camel-driver'
/ħamaamun/	'pigeons'	/ħammaamun/	'bath'
/falaaħun/	'success; victory'	/fallaaħun/	'peasant'
/qaṣaaṣun/	'punishment'	/qaṣṣaaṣun/	'story-writer'

Nevertheless, there are words in Arabic, which are of different parts of speech, that contain medial geminates for which there are no corresponding words with medial single consonants. The verbs usually take the pattern of (tafaξξala), and the nouns the pattern of (faξξaal) which normally denotes craftsmanship. For example:

/tanaffasa/	'he breathed'
/tamaħħaṣa/	'he examined carefully'
/tabaddala/	'he changed'
/takallama/	'he talked'
/taṣarrafa/	'he behaved'
/ħaddaadun/	'blacksmith'
/nadʒdʒaarun/	'carpenter'
/rassaamun/	'painter'
/dʒazzaarun/	'butcher'
/ɣawwaaṣun/	'diver'

At the same time, words having medial single consonants for which there are no corresponding words with medial geminates are too many to be counted in Arabic.

1.5.2 In I.C. Arabic

Similar to that of Classical Arabic, word-medial gemination in I.C. Arabic distinguishes one word from another, and contrasts can be made between words with geminate consonants and others with their single counterparts. In everyday Iraqi speech there exists a great number of words which are differentiated on the grounds that one contains word-medial or word-initial geminate consonant and the other contains its single cognate. With regard to I.C. Arabic this type of gemination is considered as 'geminate proper' since its presence in a word does not come as a consequence of the phenomenon of assimilation.

Word-medial geminates in I.C. Arabic are contrastive, and there are quite numerous words that contrast with others having single consonants. These words differ not only in their medial consonants, one being single and the other geminate, but also in some cases in having different vowel qualities in the first syllable of the contrasting words. This difference in vowel quality must not be considered as another factor that causes the contrasting words to yield different meanings. In such contexts, the Iraqis frequently tend to change vowel [a] of the first syllable in the Classical Arabic word into either [i] or [u]. An Iraqi layman does not differentiate between the meaning of, let's say, /kitab/ and /katab/, or /ʕubar/ and /ʕabar/ except that the pronunciation of the word with the first syllable having the vowel [a] seems to him rather pedantic and more of the literary-like style. The geminates can both be solar and lunar, as shown in the examples below:

/ħisab/	/ħasab/	'he counted'[1]	/ħassab/	'he reckoned'
/guṣab/	/gaṣab/	'reeds'	/gaṣṣab/	'he slaughtered'
/baʃar/	/baʃar/	'human beings'	/baʃʃar/	'he announced (good news)'
/ṣaxar/	/ṣaxar/	'rocks'	/ṣaxxar/	'he requested to do'
/ziξal/	/zaξal/	'he became angry'[2]	/zaξξal/	'he made (someone) angry'
/ridʒaξ/	/radʒaξ/	'he returned'	/radʒdʒaξ/	'he returned (something)'
/baṭal/	/baṭal/	'hero'	/baṭṭal/	'he gave up'
/wilad/	/walad/	'he was born'[3]	/wallad/	'he generated'
/ʃaħam/	/ʃaħam/	'fat' (n.)	/ʃaħħam/	'he became fat'
/mahad/	/mahad/	'cradle'	/mahhad/	'he prepared (the way)'

In I.C. Arabic, geminate proper is not only confined to word-medial position. It also exists initially in certain words beginning with [b], [n] or [l]. Compared with corresponding words beginning with single consonants, the words beginning with geminate consonants are used to exemplify certain grammatical structures, as in the following examples:

single		geminate	
/beeti/	'my house'	/bbeeti/	'in my house'
/baγdaad/	'Baghdad'	/bbaγdaad/	'in Baghdad'
/nisa/	'he forgot'	/nnisa/	'he was forgotten'
/libas/	'he put on his clothes'	/llibas/	'it was put on'[4]
/lizam/	'he caught'	/llizam/	'he was caught'

---

1. The pronunciation /ħasab/ may also mean 'noble birth' .
2. The pronunciation /zaξal/ may also mean the noun 'anger'.
3. The pronunciation /walad/ may also mean 'boy, lad'.
4. Metaphorically, it means 'he was neglected'.

These examples show clearly that in the case of initial [bb], the word denotes an adverbial phrase of place, whereas in the case of initial [nn] and [ll], the words indicate the passive form of a past tense.

Finally, it is of interest to remark that with some words, initial geminates are able to change an utterance from a statement into a question. This is usually the case with the following sibilants:

/ʃifit/	'I saw'	/ʃʃifit/	'what did you see?'
/sadʒdʒalit/	'I recorded'	/ssadʒdʒalit/	'what did you record?'
/ṣihit/	'I shouted'	/ṣṣihit/	'what did you shout?'

1.6 Studies on Arabic Gemination

In this section we shall review previous works that dealt with the phenomenon of gemination in Arabic as a whole. Works that studied gemination and other related subjects in languages other than Arabic will be fully reviewed in the next chapter. The section will be divided into two subsections. The first will deal with those works that studied Arabic gemination theoretically, and the second will present those works that investigated the phenomenon experimentally. The theoretical studies were, undoubtedly, greater in number than the experimental investigations since the former had started much earlier.

No attempt will be made in this section to review those studies initiated by the old Arab grammarians; this is partly because their studies lie outside the scope of the present investigation, and partly because their basic ideas about gemination in Arabic have been briefly reviewed in the first two sections of this chapter. Therefore, our remarks will be confined mostly to those studies that have appeared since the advent of the current century.

1.6.1 Theoretical Studies

During the second half of the last century two works (Forbes, 1863; Socin, 1895) were published that studied the grammar of the Arabic language. Both works discussed the Arabic term 'taʃdiid', i.e. gemination. For Forbes (ibid.) 'taʃdiid' simply signifies corroboration, and in Arabic writing a consonant is shown to be geminated when the mark (*w*) is placed over it. He states that "In certain instances the symbol taʃdiid is employed for the sake of euphony, when an inert letter is followed by a different letter having a cognate sound, or one which is capable of coalescing with the former." [p.15]. Thus, the word /madadtu/ 'I extended' is said to be preferably pronounced as /madattu/, i.e. with geminate [tt], to show that [d] has coalesced with the succeeding [t].

Forbes (ibid.) recognizes the possibility of having gemination when the [l] of the definite article /ʔal/ coalesces with a word-initial solar consonant, as in /ʔan-nuuru/ 'the light' instead of /ʔal-nuuru/. He then states that in modern Egyptian and Syrian dialects the [l] of the article "retains its natural sound before all letters, whether solar of lunar." [p.37]. He also adds that in some instances the final [n] of the two words /min/ and /ʔan/ is changed into that which follows, as in the two phrases /mim-man/ 'from whom' for /min-man/, and /ʔal-laa/ 'lest, so that not' for /ʔan-laa/.

Socin (op.cit.) follows Forbes in using the term 'taʃdiid'. According to him the doubling of a consonant, i.e. gemination, is either due "to the essential nature of the form, nominal or verbal..., or is the result of assimilation." [p.11]. He remarks that when one consonant is assimilated to another, the assimilation is further graphically represented by the removal of the 'sukuun', viz. the pausal sign, from the assimilated consonant. This, he adds, applies to the [l] of the article /ʔal/, when the latter precedes a solar consonant. The words /min, /ʕan/, /ʔan/ and /ʔin/, when followed by words beginning with [m] or [l], are said to be usually combined with them into one word and the final [n] is at the same time assimilated to the following consonant. Hence, he refers to the same words reported by Forbes (op.cit.).

Early in this century, Oussani (1901) wrote an article on what he called the 'modern' Arabic dialect of Baghdad, which he regards as one among the most important Arabic dialects. In this study he confines himself to the way consonants and vowels are pronounced and "to some rough notes on the principal phonetic, morphological and lexicographical peculiarities" that characterize the dialect. The words presented as examples clearly denote that the author was primarily concerned with that accent of Baghdadi dialect as spoken by Muslim, Christian and Jewish communities living in Baghdad at that time. However, the dialect described was "essentially Christian Baghdadi." (Blanc, 1964, p.50). For instance, words such as /dahab/ 'gold', /toob/ 'cloth, shirt', /madrasi/ 'school' and /ˀabuunuu/ 'his father' are reported in place of the most popular pronunciations /δahab/, /θoob/, /madrasa/ and /ʔabuuh/, respectively. Other examples where [r} is replaced by [γ], as in /ξaʃγa/ 'ten' and /γummaan/ 'pomegranates' are also reported. Such pronunciation can by no manner of means be considered as genuine Baghdadi dialect. This is obviously due to the fact that this type of pronunciation is not used by the vast majority of the people living in Baghdad nor in places around it. In point of fact, this is the type of pronunciation which is predominantly heard in some other places in the north of Iraq, such as in the big city of Mosul.

Although Oussani (op.cit.) does not mention the terms 'gemination' and 'geminate consonants', he quotes a number of words containing word-medial geminates. For example, the following words have been cited:

/baṭṭiix/	'melon'
/tṣalli/	'you pray'
/mdarris/	'instructor'
/mξallim/	'teacher'
/minnuu/	'from him'

At the same time, some proper nouns used in the dialect of Baghdad by the three communities, which contain geminate consonants, are also mentioned. Personal names like

/ʃakkuurii/, /razzuuqii/, /ξabbuudii/, /naṣṣuurii/, /dʒabbuurii/ and /faṭṭuuma/ are reported, in which there is a medial geminate: [kk], [zz], [bb], [ṣṣ], [bb] and [ṭṭ], respectively. A list of adverbs and adjectives of common use in the dialect is adumbrated. Some of the words included in the list have intervocalic geminate consonants, such as:

/hammeena/	'also'
/ʃwajja/	'little'
/dʒawwa/	'inside'
/barra/	'outside'

Later, van Ess (1917, 1938) published his well-known pamphlet entitled 'The Spoken Arabic of Iraq'. The widely-used Arabic term 'ʃadda' is clearly explained. It is defined as a special sign used in Arabic orthography which is normally placed over a consonant to indicate that "the consonant must be doubled." [p.5]. The most popular Arabic name 'muħammed', with medial geminate [mm], is given as an illustration.

Van Ess (op.cit.) mentions the possibility of gemination as a consequence of the process of assimilation in word-initial position. This usually happens when the definite article is followed by a solar consonant where the [l] of the article is assimilated to that consonant, which is then doubled. The word /iʃʃams/ 'the sun' is given as an example. He also remarks that "when the article is preceded by a vowel, the [i] of the article should rightly be dropped in transliteration, and the [l] affixed to the preceding word, inasmuch as its corresponding vowel in Arabic disappears in pronunciation." [p.9].

Blanc (1952, 1964) studied Arabic gemination. He states that gemination in Arabic is a prolongation of the 'continuants' and a lengthened closure of 'stops'. (Blanc, 1952, p.73). Closely following the investigation carried out by Oussani (op.cit.), Blanc (1964) examines three distinct dialects of what he called the colloquials of the Muslim, Christian and Jewish communities of Baghdad. He describes the three dialects thoroughly and compared them to one another. He also traces their historical developments from Classical Arabic.

In discussing the phonological status of consonant clusters, Blanc (ibid.) contends that the three dialects do not differ markedly in the combinations of consonants that occur as initial or medial clusters, and that a consonant tends to be assimilated to the following consonant unless that consonant is a 'sonorant' or [ʔ] or [ξ]. The three dialects show a marked tendency toward assimilation of a plosive to a following homorganic fricative. This type of assimilation is said to result in consonant gemination. Only two examples have been given to illustrate geminated consonants as a consequence of an assimilation between a plosive and a subsequent homorganic fricative. The two examples are:

[b] plus /faras/	becomes	/ffaras/	'with a horse'
[t] plus /sawwi/	becomes	/ssawwi/	'you do'

Blanc (ibid.) believes that in all three dialects word-final geminates are treated alike and that word-medial geminate clusters are reduced to a single consonant when they are immediately followed by another consonant, as in the word /salmuu/ 'greet'. This reduction, he remarks, is either due to historical or morphological processes. Word-final geminates, in the pausal form, are similarly reduced to their single counterparts so that /kil/ 'eat' and /kill/ 'all' are regarded to be homophonous in the dialects of the Christian and Jewish communities. He adds that "the distinction retained in the notation is morphophonemic, and reflects the fact that gemination is restored not only with vowel-initial suffixes ... but also when followed by a word with initial vowel, including anaptyctic vowels ... ." [p.54].

Mitchell (1956) uses the term 'doubled consonants' to mean 'geminate consonants'. For him any of the Arabic consonants may be geminated, and a geminate consonant "must be pronounced approximately twice as long as a single one." [p.8]. He suggests that special attention should be paid to the feature of gemination when it occurs at the end of words in Arabic, as in /ħabb/ 'he liked', and /muhimm/ 'important' where word-final geminate consonant is characterized by greater muscular tension in the articulatory organs.

Harrell (1957) examines consonant clustering in colloquial Egyptian Arabic. He regards it as one of the essential problems of consonant distribution in this dialect. According to him, only single consonants occur in word-initial position; while both single and abutting consonants, including geminates, occur word-medially. In word-final position single and compound consonants, excluding geminates, are said to occur.[1]

---

[1] Harrell follows Stetson's (1951) definitions of the two terms 'abutting' and 'compound' consonants. For further details see Harrell, ibid., p.30, fn.(5).

Harrell (op.cit.) disagrees with the previous analyses of Egyptian Arabic that have included final geminates. He argues that the problem mainly lies in the definition of a geminate phonetically and states that internal abutting consonants and final compound consonants are fundamentally different in their syllable functions. In abutting consonants the first consonant closes the preceding syllable and the second consonant releases the following syllable. He defines an internal geminate as "a given consonant articulation held across a syllable boundary, so that a fall of air pressure behind the closure after the formation of the articulation is followed by a rise of air pressure from a new intercostal pulse before the release of the articulation." [p.31]. In accord with this definition, Harrell comments that the crucial factor is not length per se, but "length across a syllable boundary so that a given articulation functions independently in two syllables." [loc.cit.]. On the other hand, a 'prepause geminate' is definable only in terms of length because a final compound consonant functions in only one syllable. However, Egyptian Arabic is said to show no organized use of prepausal consonant length.

Harrell (op.cit.) believes that previous analyses of prepause geminates in Egyptian Arabic suffered from a confusion of morphological alternation with phonological facts. He states that there are numerous examples of stressed prepause syllables with short vowels closed by single consonants which can stand in morphological alternation with other utterances having word-medial geminates, as, for instance, in /sikit/ 'I became silent' and /masikittiʃ/ 'I did not become silent'. He comments that this alternation does not hold consistently throughout this Arabic dialect, and suggests that "If this alternation were consistent throughout the language, it might be justifiable to treat prepause ['VC] as /VCC/." [p.31]. He, eventually, reaches the conclusion that due to lack of consistency in morphological alternation and phonetic evidence of indeterminate length of final consonants it is quite impossible to give a meaningful definition of final geminate in the Arabic dialect of Egypt.

In a subsequent paper, Harrell (1962) discusses the existence of geminate consonants in Moroccan Arabic. He distinguishes three contrasting lengths in this dialect.[1] Moroccan consonants, he remarks, not only show a contrast between short and long, but also a contrast among short, long and extra long, as it is clear in /simha/ 'her name', /sim·ha/ 'her poison' and /sim:ha/ 'he poisoned her', respectively. This type of length contrast also occurs with other consonants, such as voiceless fricatives and plosives. The long consonants are phonemically interpreted as geminates, and they are "of no particular theoretical interest." [p.696]. The extra long consonants, on the other hand, are regarded as the rare phenomenon of a regular distinction of the three degrees of consonant length in a language.

Cowan (1958) considers 'doubled', i.e. geminated, consonants in Arabic as two identical consonants that come together and which are not separated by a vowel. He states

---

[1] Cf. Lehiste (1960, 1966) and Lehiste et al. (1973) for Estonian. Also see Chapter Two of this study, section 2.2.

that vowelless dental consonants are generally assimilated to a following [t], as in /quttu/ for /qudtu/ 'I led'. The latter pronunciation is by no means considered incorrect, but the correct pronunciation of the two adjacent consonants would bring about this assimilation. In so far as writing is concerned, the first consonant is written without the sign of gemination, viz. 'ʃadda'; whereas the second receives it.

Cowan (op.cit.) also states that a vowelless [n] assimilates to a following [l] "either in pronunciation or actually in writing" [pp.5-6]; as in the conjunctions /ˀallaa/ for /ˀanlaa/ 'that not', and /ˀillaa/ for /ˀin laa/ 'if not, otherwise'. He adds that the [l] sound of the definite article in Arabic is assimilated to the following consonant when the noun or adjective defined by the article begins with one of the solar 'letters'.

In 1959, Nasr made a phonological study for predicting the location of stress in Lebanese Arabic. He presents examples in which geminate consonants occur finally in monosyllabic words, such as /ṣaff/ 'class'. Word-medial geminates are illustrated in words of disyllabic, trisyllabic and quadrisyllabic structures, as in /ʔanni/ 'that I', /sidʒdʒaadi/ 'carpet' and /sidʒdʒaadaatna/ 'our carpets', respectively.

In his article of (1960), Nasr states that length in spoken Lebanese Arabic is phonemic. It applies to consonantal and vocalic segments separately. He considers length as a segmental feature rather than a suprasegmental one. This is to say, it does not apply to syllable, morphemes of more than one phoneme, or to words as a whole. He contends that the difference between the short, i.e. non-geminate, and long, i.e. geminate, sounds is that "the long sounds take a relatively longer time to be completely produced than the short ones." [p.210]. He chooses a number of minimal pairs containing word-medial single and geminate consonants to illustrate the difference. The following are some of these pairs:[1]

/sabab/	'cause'	/sabbab/	'he caused'
/baṭal/	'hero'	/baṭṭal/	'he ceased'
/nafas/	'breath'	/naffas/	'it lost air'
/naξam/	'yes'	/naξξam/	'he smoothed'
/sama/	'sky'	/samma/	'he named'
/ξala/	'on'	/ξalla/	'he raised'
/mara/	'woman'	/marra/	'once'

Perhaps the most thorough and wide-ranging study of Iraqi Arabic ever carried out was by Erwin (1963). The study is essentially concerned with examining the grammatical features in the dialect. However, other phonological, morphological and syntactical features are tackled. For instance, geminate consonants are dealt with under the general title of 'sounds in combinations'. He views the phenomenon of gemination merely as the doubling of consonants and assumes that any consonant in Iraqi Arabic may occur double. This obviously refers to geminate consonants occurring in word-medial position irrespective of whether they are solar or lunar, as in /dabbar/ 'he arranged', /hassa/ 'now', /salla/ 'basket', /raggi/ 'watermelon', etc. He then

---

1. For more examples, see Nasr, 1960, pp.210-211.

notes that not all geminate consonants occur word-initially, and most of those geminates which exist in such word position "are the result of a combination of certain one-consonant prefixes and stems which happen to begin with the same consonant." [p.28]. The examples he presents are, in fact, either the result of assimilation between the definite article and a following consonant, or the result of certain grammatical processes. The following words are among the examples quoted:

/θθaani/	'the second'
/zzibid/	'the butter'
/ṭṭeer/	'the bird'
/ddaxxal/	'he interfered'
/mmaʃʃiṭ/	'having combed'

In word-final position, a geminate consonant is the same as the corresponding single consonant. For him the difference lies in the location of the stress. He claims that a word ending in a geminate consonant has stress on the final syllable, as in /ma'ħall/ 'place'; whereas a word ending in a non-geminate consonant has stress elsewhere, provided that the consonant should not be preceded by a long vowel, as in /'siħal/ 'he pulled'.

Erwin (op.cit.) remarks that with a geminate plosive consonant "the closure is maintained slightly longer before release than for the corresponding single consonant." [p.27]. The case seems to be similar with a geminate affricate and its single cognate. On the other hand, when a fricative, nasal, or lateral is geminated "the continuous sound is prolonged slightly longer than for the corresponding single consonant." [loc.cit.]. He, therefore, conceives of gemination as essentially the lengthening of corresponding non-geminate consonants. Geminate semivowels are reported to occur intervocalically only, as in /huwwa/ 'he' and /hijja/ 'she'.

In 1969, Erwin issued a companion volume for his work of 1963. In point of fact, it was another comprehensive study of Iraqi Arabic. Most of the basic ideas discussed in his former work are revised, elaborated and represented. Here, again, he discusses the vital significance of gemination in Iraqi Arabic and suggests that a geminate consonant "may be thought of as a special kind of two-consonant cluster in which the two are identical." [p.19]. He assumes that the distinction between a single consonant and its geminate counterpart in Iraqi Arabic is as important as distinguishing between [p] and [b] in English. This is because "the difference between the two may alone make the difference between one word and another, or between a real word and a form which will not be understood at all." [pp.19-20].

In terms of articulation, Erwin (ibid.) continues his assumption that the difference between a geminate and a non-geminate consonant is only a matter of relative duration and states that geminate consonants are sometimes called 'long consonants'. For him the difference between a single and a geminate consonant is very much like the difference between a short and a long vowel.

Cowell (1964) seems to be in agreement with Erwin's assumption. He considers the main difference between long and short sounds is simply the relative length of time the articulation is held. He states that long consonants are held not only longer but generally also tighter than short consonants. The following words, selected from Syrian Arabic, are cited to show the difference:

short		long	
/ɣani/	'rich'	/ɣanni/	'sing'
/mara/	'a woman'	/marra/	'a time'
/ħamaam/	'pigeons'	/ħammaam/	'bath'

Cowell (op.cit.) uses the two terms 'double' and 'long' synonymously. They both stand for geminate or geminated sounds. The term 'short' is used to mean single or non-geminate. He notes that in Syrian Arabic any single consonant may occur initially, medially, or finally, before or after any vowel in a word. Double consonants, on the other hand, may occur word-medially, i.e. between vowels, as in /rabbi/ 'my God', /ħatta/ 'until', /ʕammi/ 'my uncle', etc. He adds that in word-initial position, double consonants are limited to those formed by the combination of a prefix or proclitic with the first stem consonant. Consonants [f, g, ɣ, h, ħ, k, q, x, ʕ and ʔ] are excluded because they do not occur doubled word-initially. In word-final position, any double consonant may occur after a stressed vowel.

Cowell (op.cit.) remarks that at the end of a phrase "long consonants (like long vowels) do not actually contrast with short ones; writing them double simply serves to show the position of the accent and their potential significant length before vowels." [pp.23-24]. He states that long consonants seldom occur before another consonant, except in sequences involving the definite article, or demonstrative proclitics, or the person suffix-t, as in /z-zabuun/ 'the customer', /har-rʒaal/ 'these men' and /baξattna/ 'you sent us'.[1] A long consonant occurring before a vowel has the tendency to lose its distinctive length before another consonant or at the end of a phrase. Cowell also mentions the possibility of having a geminate as a result of assimilation, such as in the expressions /ʔaḥsal-lak/ 'it would be better for you' for /ʔaḥsan-lak/, and /mammuut/ 'we don't die' for /manmuut/. Assimilation that results in gemination may also occur across word boundaries, particularly between sequences of two contiguous consonants. Such assimilation tends to eliminate certain 'awkward' clusters, such as /raaξ ξal balad/ for /raaḥ ξal balad/ 'he went to town'.

In his study of the eastern Arabian dialects, Johnstone (1967) reports that word-initial geminate consonants are not very common in these dialects, except in very few phrases such as /ssalaam/ 'the peace' and /ttamr/ 'the dates'. These geminate consonants may change into their single counterparts, as in the everyday greeting /salaam ξaleekum/ 'peace be on you' for /ssalaam ξaleekum/. Johnstone shares Mitchell's view (1956) that a final geminate is characterized by tenseness of articulation as compared with a non-geminate. He

---

[1] In the case of /z-zabuun/, it is difficult to see on what evidence his conclusion is based, i.e. that the first [z] is considered as a long consonant. Obviously, it occurs as a result of the process of assimilation between the definite article and the following [z].

states that gemination after a short vowel may exist as a natural consequence of 'stress regression'. This is illustrated by pronouncing the word /balamm/ 'a type of boat' with primary stress on the second syllable /-lamm/ and with a rising intonation as if requesting repetition or confirmation of information.[1]

In eastern Arabian dialects, 'deliberate' speech is characterized by frequent occurrences of consonant clusters due to elision of short vowels. One element of such consonant clusters is a geminate consonant which usually becomes a single consonant. For example:[2]

/jdʒassimuun/	>	/jdʒassmuun/	>	/jdʒasmuun/	'they distribute'
/dʒassimih/	>	/dʒassmih/	>	/dʒasmih/	'distribute it'

In these cases, however, the geminate may be retained when it is a 'liquid'. It may also be retained in word junction, as in /kul(l) joom/ 'everyday'.

Al-Ani and May (1973) state that there exists a phonemic as well as a lexical contrast between pairs of words differentiated merely by gemination of the second consonant of the root. Pairs of words, such as /kataba/ 'he wrote' and /kattaba/ 'he made (someone) write'; /kasara/ 'he broke' and /kassara/ 'he broke (it) into pieces' are reported. They comment that despite the substantial role played by acoustic phonetics in demonstrating that "a sequence of two identical consonants is indivisible into two units,

---

[1] Perceptually, the consequence is that the word sounds like */-llamm/.

[2] See Johnstone, op.cit., p.27.

there is also contrast in the articulation of such geminate/ non-geminate pairs." [p.121]. They contend that the geminate consonants of /kattaba/ and /kassara/ cannot be physically considered as sequences of [tt] and [ss], respectively. This is believed to be due to the fact that, for instance, a geminate plosive like [tt] shares with its single counterpart [t] the phonetic phenomenon of one release, i.e. occlusion offset, rather than two, in spite of the fact that the geminate [tt] is held for a longer closure duration than [t]. This principle applies similarly to fricative consonants where the difference of duration between single consonants and their geminate cognates is clearly noticeable.

While discussing the syllabic structure in Arabic, Al-Ani and May (op.cit.) remark that, on the phonetic and not on the phonotactic level, one encounters the problem of segmenting clusters of two identical consonants when they occur word-medially. According to them, the indissolubility of geminate consonant clusters is a phonetic phenomenon which creates a difficulty in segmenting the syllable in Arabic on the phonetic level. They state that the difficulty lies in the fact that a syllable boundary signal should occur between the geminated consonants, and that a geminate consonant in Arabic is represented by two identical symbols in the transcription only for the sake of consistency with the phonological system of the language.

In his two recent studies of the phenomenon of consonant gemination in Arabic, Benhallam (1979, 1980) tries to establish rules based on the generative approach to phonology. He reports that the analysis of geminate clusters shows that one needs to go beyond stating that a rule has to be constrained from breaking up geminate clusters is more highly valued than one that breaks them. He distinguishes between two types of geminates in Arabic; the 'underlying geminates' and the 'derived geminates'. The former can be split up by morphological or phonolexical rules; whereas the latter can be split up only by phonological rules. He states that geminate clusters can be split up by morphological or morpholexical rules but not by phonological rules. Unfortunately, Benhallam does not present examples to illustrate this statement, i.e. what he calls 'the geminate law'. He only remarks that this law does not claim to be universal and that "it is a claim as to a general tendency of how geminate clusters behave in natural language." (Benhallam, 1980, p.141). He points out that counterevidence to this law can be constituted by a case of geminates that are broken up by phonological rules, as in the following words from Moroccan Arabic:

mamdud	(>/m+mdd/)	'lying down'
mamluk	(>/m+mlk/)	'owned'

Both words are obviously passive participles and the resultant word-initial geminates are formed by affixation. Benhallam states that an investigation into the use of the above words reveals that they are heard in 'very careful speech' where they can be considered as cases of hypercorrection. In 'normal speech', however, one usually hears <u>mmdud</u> and <u>mmluk</u>, respectively.[1]

---

[1] For further details see Benhallam, 1980, Chapter Seven, pp.125-147.

Despite the fact that both studies are mainly concerned with Moroccan Arabic, Benhallam extracts his rules from words which are frequently used by speakers of modern standard Arabic. He contends that triconsonantal root is the preferred type in Arabic because it is 'the minimal viable pattern'. He adds that "if a great number of biconsonantal roots are allowed into the language and a process of weakening or consonant elision takes place, one would end up with monoconsonantals." (Benhallam, op. cit., p.5). However, for him all monosyllabics in Arabic are considered as functional words with a very minimal semantic load. They are mostly prepositions which have to be connected to the following word because they cannot stand by themselves. On the other hand, he remarks that triconsonantals are not allowed to have initial geminates "so that the geminates would not undergo a likely simplification (shortening ?) and end up as biconsonantals." [loc.cit.].

Following Kurylowicz (1972), Banhallam (op.cit.) believes that triconsonantals can be created from original biconsonantals. For instance, the doubling of the second consonant creates verbs with word-final geminates, as in the examples below:

γ-l	γ-l-l	/γalla/	'he inserted'
ʃ-q	ʃ-q-q	/ʃaqqa/	'he split'
k-l	k-l-l	/kalla/	'he got tired of'
b-l	b-l-l	/balla/	'he moistened'

Quite obviously, Benhallam considers the above words as triconsonantal verbs ending with geminates. He disregards the presence of word-final [a] since its prime function in a word is to denote the non-pausal form in Arabic. The pausal forms, viz. /ɣall, ʃaqq, kall and ball/, respectively, are used in Arabic as well. However, he notes that not all triconsonantals with geminates are of biconsonantal origin because he thinks that "it is hard to tell just from looking at a verb whether it is of biconsonantal origin or not." [p.6]. He adds that the geminate consonants might have been there all along, or might have resulted from the process of assimilation of two homorganic 'radicals' or an expanded second radical, as is evident from the examples quoted above.

What is more, Benhallam (op.cit.) observes that Arabic has, among its enormous lexical entries, quadriconsonantal forms containing word-medial geminates which are originally derived from triconsonantal forms. The examples below illustrate this fact:

/dʒammada/	'he froze'	from	/dʒamada/	'to freeze'
/dʒaddala/	'he tightened'	from	/dʒadala/	'to tighten'
/faqqaˀa/	'he burst'	from	/faqaˀa/	'to burst'

He comments that the geminates in the above examples are the result of regressive assimilation where the original quadriconsonantal forms are /dʒalmada/, /dʒandala/ and /farqaˀa/, respectively. Both /dʒamada/ and /dʒalmada/ are regarded as synonymous; whereas /dʒammada/ is considered as the 'causative'. Nevertheless, it is hard to prove that /dʒammada/, /dʒaddala/ and /faqqaˀa/ are resulted from regressive assimilation in the quadriconsonantals /dʒalmada/, /dʒandala/ and /farqaˀa/, respectively. This is partly because these words are derived from different roots, and partly because they bear different semantic interpretations. Accordingly, Benhallam's suggestion that one should be cautious not to confuse forms that have doubled their medial consonant, as in /kattaba/, and forms that result from regressive assimilation, as in /dʒammada/, seems irrelevant since the geminates in both forms are primarily derived from words having a triconsonantal pattern, i.e. /kataba/ and /dʒamada/, respectively, where the medial consonant is single.

More recently, Rahim (1980) has made an interesting and comprehensive investigation of I.C. Arabic. He studies the phonology of everyday Iraqi Arabic from the functional point of view, and attempts, for the first time, to arrive at a 'phonematic' analysis of I.C. Arabic entirely based on the theories established by the functional school of linguistics.

Following other researchers of this school (e.g. Cantineau, 1960a and 1960b), Rahim (ibid.) prefers to use the widely-used Arabic term 'taʃdiid' and its derivatives in lieu of the European term 'gemination' and its derivatives.[1] He remarks that 'taʃdiid' is a term very often used by the Arab grammarians and it has never been used by non-Arab linguists; while the term 'gemination' is generally used by European and American scholars. He adds that "the concept of 'gemination', on the one hand, mostly refers to the 'doubling of consonants'. On the other hand, the term 'tashdīd'

---

[1] For full information see Rahim, ibid., p.196, fn.2.

literally means 'reinforcement' while in Arabic linguistic terminology it means more or less the same as 'gemination'." [p.196].[1]

In explaining his own views of the phenomenon of gemination and its phonological status, Rahim (op.cit.) seems to follow closely Erwin's (1963, 1969) interpretations of the same phonetic feature. Phonetically, he considers gemination in I.C. Arabic simply as the lengthening of single consonants so as to yield geminate consonants. He shares Erwins views in that when plosive and affricate consonants occur geminated, "the closure is maintained longer before the release stage of the corresponding 'single' stops." [pp.203-204].[2] On the other hand, when fricative consonants occur geminated, the friction noise is prolonged slightly more than the corresponding single fricatives. The nasals and laterals are said to behave similarly, and with a geminate flap the tapping of the tongue blade is made to recur several times.

Rahim agrees with previous views on gemination, especially those reported by Cantineau (op.cit.) and Erwin (op.cit.) in that gemination occurs where a consonant cluster is likely to occur. He assumes that consonant clusters in I.C. Arabic may occur word-initially, word-medially and, less frequently, word-finally. These can be shown in the examples below, respectively:

---

[1] Underlining mine.

[2] Rahim (op.cit.) uses the term 'stop consonants' to represent both affricates and plosives.

/snuun/	'teeth'
/ħafla/	'party'
/ʃarq/	'east'

Consequently, gemination in I.C. Arabic is likely to occur in identical word positions.

On the other hand, in discussing the phonological status of gemination occurring in different word positions in I.C. Arabic, Rahim (op.cit.) states that both in word-initial and word-medial positions, gemination functions distinctively as there exist in I.C. Arabic plenty of minimal pairs which are merely based on the distinction between geminate and non-geminate consonants. He also mentions the possibility of having gemination in word-initial position as a consequence of assimilation between the definite article and an ensuing consonant. While all consonants in I.C. Arabic can occur geminated in word-medial position, only some of them are reported to occur in word-initial position. This essentially depends on the type of consonant occurring in this position and on whether it is liable to assimilate with the preceding definite article.

Rahim (op.cit.) claims that gemination in word-final position can, in very rare cases, function distinctively. Unlike that in word-initial and word-medial positions, gemination occurring in word-final position is "not straightforward", and it is regarded as non-distinctive since it is only a "phonetic lengthening" of I.C. Arabic consonants. In this respect, Rahim gives more details when he states that "This is so only in those SIA words where one of the respective identical consonants is commutable with other SIA consonants so as to produce words that belong to CA but which are retained in the speech of most educated speakers of SIA. Such words would always be monosyllables with a short vowel." [p.207].[1] He, then, briefly concludes that gemination in word-final position affects most I.C. Arabic consonants distinctively in very few cases and non-distinctively in most cases.

1.6.2 Experimental Studies

In 1957, Mitchell made an experimental study of gemination in two dialects; the Bedouin Arabic of the Cyrenaican Jebel and the Berber of Zuara. Two instrumental techniques have been exploited in this study; the palatograph and the kymograph. The term 'length' is equated both with the phonological term 'quantity' as well as with the phonetic term 'duration'.

In Cyrenaican Arabic, Mitchell (ibid.) investigates the article-noun forms of a number of different patterns, forms of the verb, forms with third person pronominal suffixes, and consonant-length and vowel-length. He sets up a 'commutation system' of three prosodies: geminate cluster (gc), non-geminate cluster ($\overline{gc}$), and non-cluster ($\bar{c}$). Sub-systems, each of two commutable terms, are established for the last two prosodies, namely non-geminate cluster ($\overline{gc}$) and non-cluster ($\bar{c}$). The categories set up are totally phonological and are never established on purely phonetic grounds.

---

[1] SIA stands for 'Spoken Iraqi Arabic' and CA for 'Classical Arabic'.

Mitchell (op.cit.) remarks that each verb form in Cyrenaican Arabic is characterized by a given affix as well as by "an associated scatter of cognate forms (verbal noun and participles)." [p.186]. Examples such as /ka'sar/ 'he broke' versus /'kassar/ 'he smashed', /ga'sam/ 'he divided (in two)' versus /'gassam/ 'he divided up', and /ξa'dam/ 'he executed' versus /'ξaddam/ 'he destroyed' exhibit the following relevant phonetic features:

(i) length of the medial consonant
(ii) tensity of the medial consonant
(iii) stress-incidence
(iv) quality of the penulatimate vowel

These features clearly denote that non-geminate medial consonants are always associated with pre-stressed position; while, on the other hand, the geminates are always associated with post-stressed position and open vowel quality in the penultimate syllable. He regards these features as "co-exponents respectively of non-cluster and geminate cluster considered as commutable prosodic categories." [loc.cit.].

Mitchell (op.cit.) considers the palatograms and kymograms obtained from investigating the Berber dialect of Zuara of general interest. They not only show "the impossibility of defining the prosodic categories of cluster, both geminate and non-geminate, and non-cluster a priori on the basis of length of articulation and/or any other phonetic feature, but also provide some additional justification for a prosodic approach which perceptual observation suggests."

[p.193]. He finally summarizes the aspects of the prosodic view of gemination and concludes that what may be called relative phonetic length may often be included among the exponent of gemination, and it is impossible to reverse the procedure and consider phonetic length as the criterion for the category.

Obrecht (1965) conducted three experiments in which synthetic stimuli are employed to investigate duration as a cue factor in the discrimination of single versus geminate consonants in Arabic. Contrasts in three phonetic categories, viz. 'stops' [b] versus [bb], 'nasals' [n] versus [nn], and 'spirants' [ṣ] versus [ṣṣ], are involved in this study. Both the stop and the nasal contrasts are used in word-medial position; whereas the spirant contrast is found in word-initial position. Native speakers of Arabic, from different Arab countries, participate as subjects in this investigation. Spectrographic observations are made of the selected word pairs, and from these a number of variations of each are synthesized by using the well-known 'Pattern-Playback' at the Haskins Laboratories.

In order to locate the perceptual boundary between [b] and [bb], a series of seven stimuli were synthesized "which varied only in the duration of the silent period representing the duration of stop closure." [p.32]. The words studied are /xabar/ 'news' and /xabbar/ 'he informed' with primary stress on their first syllables. At the same time, the minimal pair /bana/ 'he built' and /banna/ 'mason' are chosen for another experiment designed for the purpose of locating the perceptual boundary between intervocalic [n] and [nn], respectively.

In the case of designing a discrimination test to investigate the perceptual boundary between single and geminate spirant consonants, Obrecht (op.cit.) uses the minimal pair /ṣabij/ 'boy' and /ṣṣabij/ 'the boy' as the experimental frame. He finds this pair representing an interesting grammatical contrast besides its contrast on the phonological level as single versus geminate. He states that "this particular pair is apparently found only in the Arabic of the eastern Mediterranean area." [p.37].

The final results of the three experiments show that the values found for the perceptual boundaries are of interest. Summarizing, Obrecht (op.cit.) reports that the theoretical boundary point, or the point of maximal confusion, appeared in the three experiments are as follows:

(a) For the stops, [b] versus [bb], 140-160 msec;
(b) For the nasals, [n] versus [nn], 90-110 msec;
(c) For the spirants, [ṣ] versus [ṣṣ], 140-160 msec.

The results also indicate that clear identification boundaries are found between the contrasts in the three phonetic classes which give an indication of the degree of exclusivity of duration as the determining cue factor in distinguishing between a single and a geminate consonant in Arabic. The results also show that "the structural status of gemination influences decisively the sharpness of discrimination." [p.31].

In 1966, Al-Ani started an acoustical and physiological investigation of the phonology of Arabic. This work was later published in a book in 1970. The acoustical aspect of the investigation was mainly based on spectrographic analysis; while the physiological aspect depended heavily on examination of X-ray films.

Al-Ani (1970) states that "The relative duration of the consonants depends upon whether they occur initially, medially or finally." [p.75]. He adds that consonant duration also depends on whether the consonants are aspirated or non-aspirated, voiced or voiceless, and single or geminated. According to him gemination involves prolongation of the 'continuants' and a longer closure of 'stops'. Here he follows Blanc's definition of the phenomenon of gemination reported in his work of 1952. Geminated consonants and long consonants are regarded as being synonymous and treated as identical clusters.

Al-Ani (op.cit.) believes that every consonant cluster, whether its components are identical or non-identical, involves "a close transition which means that the first member of the cluster - always occurring as a coda - is not released until the second member - always occurring as an onset - is uttered." [p.77]. In as far as the syllable boundary in Arabic is concerned, he states that the first member of a consonant cluster occurs as the coda, i.e. offset, of the preceding syllable, and the second member always occurs as an onset of the following syllable. He agrees with other scholars that geminate consonants contrast with their corresponding single consonants. Only four words arranged in two minimal pairs have been reported to show this contrast. The words are:

/qatala/	'he killed'
/qattala/	'he slaughtered'
/kasara/	'be broke'
/kassara/	'he destroyed'

Following his previous studies of Arabic phonology, Al-Ani (1967) examined the acoustical and physiological properties of the Arabic pharyngeal consonant [ξ]. The examination was largely dependent on the speech of male informants who were native speakers of Iraqi Arabic. The consonant [ξ] was studied acoustically by means of spectrograms and was examined physiologically with the help of X-ray sound films.

Spectrograms containing word-medial [ξ] were compared with those containing word-medial [ħ] and word-medial [?]. Al-Ani (ibid.) notes that "The characteristics of the medial [ξ] are determined by their environment and whether the [ξ] is geminated or single." [p.89]. It has been found that when [ξ]is geminated, it is always a voiceless stop appearing as a silent gap, with relative duration approximately from 300-330 msec. On the other hand, word-medial geminated [ħ] is considered as a voiceless fricative appearing as noise with higher frequencies than those of word-medial geminated [ξ]. The contrast between these two consonants occurring intervocalically led the researcher to conclude that "there is a difference between the medial geminated [ξ] and the medial geminated [ħ] - the contrast being not so much that one is voiced and the other voiceless but that one is a stop and the other is a fricative." [p.90].

Spectrographic analysis also shows that word-medial geminated [ξ] appears to be similar to word-medial geminated [?], except for the neighbouring vowel formants where F2 is lower when the vowel is next to [ξ] than when it is next to [?]. This is often the case with short vowel [a] and its long counterpart [aa].

Odisho (1973) produced a very stimulating study of the 'mufaxxama' consonants in Baghdadi Arabic. He follows Abercrombie (1967) in distinguishing between the two terms 'consonant cluster' and 'abutting consonants'. The first is used to denote a sequence of more than one consonant which is restricted to one syllable; whereas the second is defined as a sequence of consonants that spreads over two syllables. Elsewhere, Odisho (1975) draws our attention to the fact that "any confusion between consonant clusters and abutting consonants, for instance treating them as being structurally and phonologically identical, will lead to a drastic difference in the phonological description of a given language. The difference will be reflected mainly in the size of the permissible clusters as well as in the structural places where they are permitted." [p.108].

Odisho (1973) rejects Erwin's statement (1963) saying that Iraqi Arabic allows a consonant cluster comprising more than two consonants in word-medial position. He claims that Erwin "does not differentiate between a consonant cluster and abutting consonants." [p.145]. He states that Arabic, no matter whether it is classical or dialectal, does not allow the existence of consonant clusters of any kind in word-medial position. What we really have is abutting consonants and not consonant cluster since the first member of the cluster belongs to one syllable and the second member to the other. Consequently, Odisho (op.cit.) suggests that cases of geminate consonant clusters do not occur word-medially in Baghdadi Arabic. They only occur word-initially as a result of the process of assimilation between the definite article and the consonant initiating the following word. These examples have been presented:

/ʔal + θaaliθ/ 'the third' becomes /θθaaliθ/

/ʔal + raabiʕ/ 'the fourth' becomes /rraabiʕ/

Odisho (op.cit.) does not mention the possibility of having geminate consonant clusters in word-final position in Iraqi Arabic. However, some geminate consonants used word-finally have been reported under words having the syllabic structure of either (CVCC) or (CCVCC) where the final CC-element frequently forms a geminate consonant, as in /farr/ 'he fled' and /ṣfarr/ 'it became yellow', respectively.[1] His failure to recognize word-final geminates is probably due to the fact that he does not consider word-final CC-element, where the (-CC) represent identical consonants, as a geminate consonant cluster in Baghdadi dialect.

Quite recently, Hassan (1981) has experimentally studied vowel duration in Iraqi spoken Arabic. The study aims at investigating some acoustic, articulatory and aero-

---

[1] More examples can be consulted in Odisho, 1973, p.147.

dynamic correlates of vowel duration, and at answering whether or not one can regard the factors governing the systematic variations of vowel duration as phonetic-universal or language-specific phenomena. The influence of intervocalic geminate consonants on vowel duration has been first acoustically examined. A table of two sections, containing items with word-medial plosives and/or fricatives, has been presented. (See, Hassan, op.cit., p.245). The first section tackles the influence of voiced and voiceless plosives as well as voiceless emphatic plosives on the preceding vowel duration in the following cases:

(a) when the vowel is in a stressed syllable, as in /'fatiħ/ 'opening' versus /'fattiħ/ 'cause to open';

(b) when the vowel is in an unstressed syllable, as in /fa'taat/ 'a girl' versus /fat'taat/ 'person or instrument that breaks up solid things into fragments'.

The second section, on the other hand, tackles the influence of intervocalic fricatives on the preceding vowel duration. Unlike section one, this section merely includes voiceless fricatives and only one pair of contrasting words has been reported, namely /'baʃar/ 'human beings' versus /'baʃʃar/ 'gave good news'. The acoustic duration values of vowel [a] in the first syllables of the above words are shown to be always longer when this vowel occurs before a single consonant than before a corresponding geminate one. The differences are said to be statistically 'very significant'. For instance, the average acoustic duration of vowel [a] in the word

/'fatiħ/ is 60 msec; while it is only 50 msec in /'fattiħ/. In the case of fricatives, [a] has the average value of 80 msec in /'baʃar/ and 65 msec in /'baʃʃar/.

Hassan (op.cit.) remarks that, no matter whether a geminate consonant in Iraqi Arabic is considered as a combination of two identical consonants or one long consonant, his final results are found to be in agreement with those results reported by Delattre (1962) in that "a vowel is shorter before a consonant cluster or a geminate consonant than before a single consonant." [p.246]. He adds that his data bear testimony to the fact reported by other researchers (e.g. Nooteboom and Slis, 1972; Thananjayarajansingham, 1976) that long consonants are usually preceded by shorter vowels. At the same time, his examination of the articulatory and aerodynamic correlates of vowel duration shows that the vowel preceding word-medial geminate [tt] in /fat'taat/ has a shorter articulatory duration and a shorter duration of vocal folds adduction; whereas the vowel preceding the non-geminate counterpart [t] in /fa'taat/ has a longer articulatory duration and a longer duration of vocal folds adduction. The geminate [tt] is also characterized by having higher intraoral pressure and longer duration of articulatory occlusion. Conversely, the non-geminate [t] is characterized by having lower intraoral pressure and shorter duration of articulatory occlusion. Moreover, the examination shows that both the first syllable and the word overall articulatory timings are very significantly longer for /fat'taat/ than for /fa'taat/.

With regard to the articulatory and aerodynamic measurements in words containing medial geminates versus corresponding words with non-geminates are concerned, Hassan (op.cit.) only depends on one minimal pair of words, viz. /fat'taat/ versus /fa'taat/. It is obvious that the first word of the pair contains a geminate plosive, and the other contains its single cognate. Neither fricatives nor voiceless emphatic plosives are investigated here. It would be a good idea if the researcher extended his investigation to include the same stimulus items reported in his acoustic investigation to show the effect exerted by intervocalic geminates on vowel duration while studying the articulatory and aerodynamic correlates of vowel duration preceding geminate and non-geminate consonants as well. This will, most certainly, allow the reader to compare instantaneously the acoustic influence as well as the articulatory and aerodynamic influences of word-medial geminates on preceding vowel duration of the same word. At the same time, Hassan (op.cit.) seems unaware of the fact that the syllabic structure of the word /fat'taat/ is of the pattern (CVC + CVVC), and that of /fa'taat/ is (CV + CVVC). Phonologically speaking, he considers the syllabic structure of the first word as if it were of the pattern (CVCC + VVC) and that of the second as if it were (CVC + VVC). This is, of course, not permissible in Arabic at all because a syllable in Arabic must always begin with a consonant sound.

While discussing the 'myodynamic timings' and 'myodynamic durations' of the first syllable in both words, Hassan (op.cit.) talks about the occlusion onset and occlusion offset of the final plosive of the first syllable.[1] The average value of the myodynamic duration of the final plosive occlusion of the first syllable is reported to be 140 msec for /fat'taat/ and 55 msec for /fa'taat/. Values of peak Po and the acoustic duration of the same final consonantal segment are measured similarly. However, unlike results reported in previous studies (e.g. Stetson, 1951), Hassan concludes that his studied data show that "neither from the acoustic nor from the myodynamic and aerodynamic points of view is there any indication that gemination involves a rearticulation of the same consonant." [p.396]. He adds that they show steady and uninterrupted traces for the occlusion time of the geminate plosives. These findings are found to be compatible with those reported by other scholars (e.g. Blanc, 1952; Nasr, 1960; Al-Ani and May, 1973) in that a geminate in Arabic is articulated as one long indissoluble consonant.

---

[1] Underlining mine.

CHAPTER TWO

GENERAL SURVEY OF PREVIOUS STUDIES ON GEMINATION AND PRODUCTION OF LONG VERSUS SHORT CONSONANT DIFFERENCES

2.1 Introduction

Duration is probably one of the most extensively investigated features of speech sounds, and it is perhaps one of the phenomena in phonetics in which researchers have been most frequently interested. Studies of a number of different languages have shown that just as vowels can be long, when compared with other vowels occurring in the same phonetic context, so consonants can be long.

The question of whether there is any phonetic difference between long consonants and geminate consonants has been debated for years.[1] Scholars seem to hold different opinions in this respect. It is the purpose of this Chapter to define the term 'gemination' and review previous related studies, whether theoretical or experimental, and expose their main views on the phenomenon of gemination. Our examples will be mainly restricted to languages other than

---

[1] Distinction should be made between the two words 'geminate' and 'triple'. Both words are of Latin origin; the former means 'twin' or 'binate' and the latter stands for something 'having three parts'. Therefore, it is not appropriate to equate the term 'geminate' (see section 2.2 and footnote 1 below) with the term 'triple' or 'triplicate' since 'geminate' stands for a doubled consonant; whereas 'triple' or 'triplicate' implies three consonants, as in Estonian (for specimen, see Principles of the IPA). In consequence, one can use the term 'gemination' in parallel with the term 'triplication', where the first stands for the doubling of a consonant and the second for the tripling of a consonant.

Arabic, for the latter has been thoroughly dealt with in the first chapter of the present study. Nevertheless, Arabic will be referred to in cases where it is compared with some other languages.

2.2 Long, Double and Geminate Consonants - Phonological and Phonetic Interpretations

The term 'gemination' is usually defined[1] in terms of sequences of two identical articulations and a prolongation of the articulatory posture (Stetson, 1951; Jones, 1956, 1960; Hall, 1964; Gimson, 1970; Kenstowicz and Pyle, 1973; O'Connor, 1973; Catford, 1977). It is quite evident that this definition embraces long as well as geminate consonants. This is why for some scholars the two terms 'long' and 'geminate' are equated. Robins (1964), for instance, states that "consonants can be long, or geminate, when the closure or obstruction is held momentarily before release." [p.80] Examples have been given from English in words such as night-time and solely, as well as from Italian where consonant gemination distinguishes words such as fato 'fate' and fatto 'fact'.

By many other scholars, however, long consonants and geminate consonants are differentiated on the grounds

---

[1] The word 'geminate' is lexically defined as "a doubled consonant", and the noun 'gemination' as "the doubling of an originally single consonant" or "the doubling of a letter in the orthography". (The Oxford English Dictionary, 1933, vol. 4, p.98). Linguistically the term 'geminate' is defined as "a sequence of identical adjacent segments of a sound in a single morpheme". (Crystal, 1980, p.158). The phenomenon of 'gemination' is simply defined as "doubling or prolonging, especially of consonant sounds". (Pei, 1966, p.102). Also see Pei and Gaynor, 1954, p.81.

that geminates are characterized by having their articulation extending over two syllables. Hence, the term 'double consonants' is often used to mean 'geminate consonants'. Abercrombie (1967) refers to this fact and remarks that "double consonants also must be distinguished from long consonants. A double consonant is one whose duration extends over two syllables, whereas the duration of a long consonant is confined to a single syllable." [p.82]. In regard to English, double consonants are frequently found at word junctions, as in <u>wholly</u>, <u>unknown</u>, <u>book-case</u>, etc.

Other scholars go even further to make a distinction between geminate sequences and geminates proper, i.e. true geminates. For instance, according to Catford (op.cit) <u>the term geminate is applied only to those cases where the sequence occurs within one and the same morpheme</u>.[1] Italian is one of the best known languages in which geminates, in this sense, are very common, as in the two Italian words <u>quello</u> and <u>notte</u>. Mortimer (1977) comments on that. He says that if we accept this definition, English geminate sequences "would not be classified as geminates, since they occur only across word boundaries or in words such as unknown or guileless, containing two morphemes." [p.2]. He, therefore, suggests that in his study consonantal sequences such as occur in <u>nice cider</u>, <u>cool lager</u> and <u>black coffee</u> will be termed 'geminate consonants'. He considers the term 'double consonant' as synonymous with the term 'geminate consonant' only when he refers to sources who prefer it.

---

1. Underlining mine.

Catford (op.cit.) draws our attention to the fact that the term 'geminate' has been commonly used in place of what he calls the 'strong' or 'tense' consonants existing in east Caucasian languages. In these languages when strong and long consonants occur between vowels they may be divided between two syllables, and so they are regarded as geminate in the sense discussed by Catford. However, when these consonants occur word-initially, and hence entirely within one and the same syllable, "the geminates are merely long, and, in most east Caucasian languages, unaspirated, as opposed to the short and aspirated 'non-geminates'." [p.211]

Lehiste (1970) considers long consonants as geminate consonants and assumes that in Finnish, for example, intervocalic consonants are regarded as true geminates and normally contain a syllable boundary. She remarks that "if a language has consonant clusters that function in the same manner as long consonants, it may be useful to analyse these long consonants as clusters of identical consonants regardless of whether it is possible to demonstrate, phonetically, their geminate nature." [p.44] Again, reference is made to Finnish in which short and long consonants contrast in every position except word-initial and word-final where no clusters in native words occur in these positions. Nevertheless, long consonants continue to be treated as clusters because of being phonetically analogous.

Lehiste (ibid.) contends that the treatment of long sounds as clusters of two identical sounds "presupposes a two-way quantity opposition." [p.45]. She refers to the existence of some languages in which more than two contrastive durations can be observed. Estonian, for instance, is characterized by having over long sounds which are, in accordance with the principles of the I.P.A., written with three phonetic symbols. The representation of using three phonetic symbols to transcribe such sounds "implies an analysis of overlong sounds as clusters of three identical sounds." [p.46] Scholars seem to be in disagreement as to the question of whether Estonian has two or three vowel and consonant distinctive qualities.[1]

The appropriateness of treating long sounds as clusters of two identical sounds in many languages is widely accepted by a number of linguists (Swadesh, 1937; Gleason, 1955; Hockett, 1955; Hall, 1964; Rodon, 1967). Gleason (ibid.) believes that long consonants, in terms of their occurrence, have the same restrictions as do clusters of two consonants. This is to say, clusters of two consonants occur commonly only after short vowels, no clusters of three or more consonants, and no clusters of a long consonant and any other consonant. According to Hall (ibid., p.99) long consonants are usually to be interpreted as geminate consonants. They involve a two-consonant cluster consisting of the consonant followed by itself. Examples, in the form of minimal pairs, are quoted from Italian in which contrast between short and long consonants is shown to be significant, as in <u>carro</u> 'cart' versus <u>caro</u> 'dear'.

In his definition of long consonants, Swadesh (ibid.) includes "any consonantal sound involving the maintenance

---

[1] More details can be referred to in Lehiste, op. cit., p.46, and also in Jones, 1967, pp.119-121 and pp.132-133.

of essentially the same articulatory position and mechanism for a relatively protracted time, being long by reference to other sounds in the same language." [p.2]. For him long consonants have different phonemic status in different phonetic totalities.[1] Within his classification of 'phonetic types', Swadesh talks about 'sequences of like phonemes' in which he includes geminate clusters as sequences of identical phonemes. He also talks about sequences of homorganic, but not identical, clusters. He considers long consonants as geminate clusters, and examples have been quoted from English. For example, long 'ambisyllabic' consonants are found at the juncture of two words, e.g. shoot tigers, or of a word and certain proclitics and enclitics, e.g. unknown, saneness; but never within the word proper. The examples of long consonants are in contrast with two phonetically different kinds of short consonants occurring in the same positions, e.g. shoot elephants, and see tigers. The first example exhibits word-final consonant followed by vowel of the following word; while the second exhibits word-initial consonant preceded by final vowel of the preceding word. Phonetically speaking, Swadesh regards the long consonants to be essentially like a word-final consonant followed by a word-initial consonant. He states that "the conditions of the particular phenomenon and of the phonetics as a whole are met, if we interpret the long consonants as clusters of two identical consonants, in every other respect comparable to non-geminate clusters in the same position, e.g. - kt - in (cook tigers)." [pp.4-5].

---

[1] The meaning of the term 'totalities' is explained by Swadesh, op.cit., p.1, fn. 2.

Swadesh (op.cit.) points to the fact that ambisyllabic long consonants in the juncture of syllables within the word occur in Finnish, e.g. *europpa* 'Europe'. These ambisyllabic consonants which are conceived as geminate clusters "come within regular classes of clusters (e.g. stop + stop, nasal + nasal)." [p.5]. He comments that the pronunciation is primarily the same as that of other clusters, and this may be regarded as a further indication to interpret that the ambisyllabic long consonants are geminate clusters. However, he adds that geminate clusters pronounced as long consonants are not necessarily word-medial and ambisyllabic. Illustrating examples have been given from Hungarian and what he called 'modern Arabic' in which long consonants contrasting with short ones and comparable to other consonant clusters are shown to occur at the end of words, as in the two Hungarian words: *jobb* 'better' versus *žeb* 'pocket'. In Italian, on the other hand, long consonants which are conceived as geminate clusters "involve types of clusters which do not occur otherwise." [p.6]. Ambisyllabic long consonants occur either between vowels or between a vowel and a liquid, e.g. *mattina* 'morning' versus *matita* 'pencil', and *spettro* 'spectre' versus *teatro* 'theatre'.

Swadesh (op.cit.) also presents examples from English which are compared to others from Italian in which long consonants might be non-geminate clusters. For instance, the combination of English -d dʒ- in *good job* is pronounced "either with a characteristic d closure followed by a shift of articulating position to the dʒ position, or with a single long closure in the dʒ position." [p.7]. In these two pronunciations, the latter seems to be similar to that found in the Italian word /ɔdʒdʒi/ 'today'. He argues that the English pattern differs from that of the Italian in a number of ways: "Firstly, there is a considerable amount of accommodation of phonemes to surrounding ones in English; a d pronounced in the dʒ position before a dʒ in English is quite in keeping with the normal tendency. Secondly, English clusters are not limited in the way that clusters are limited in Italian. Finally, -d dʒ- in English contrasts with the doubly released cluster in nədʒ dʒan (nudge John); the latter cluster is clearly -dʒ dʒ- and so the former can be only -d dʒ-." [pp.7-8].

Other examples have been also reported from other different languages, such as Muskogee, Sanskrit and Shawnee.

Jones (1967) does not use the term 'gemination' or 'geminate' sounds. He distinguishes between sounds characterized with 'true' or 'indivisible' length and others which are merely 'doubling' or 'repetitions' of sounds. He thinks that it is often difficult to perceive the acoustic difference between a 'long' sound and a 'doubled' sound. By 'true' or 'indivisible' length is meant a long duration "which is not felt by the speaker as a repetition of a sound, and which cannot be replaced by doubling." [p.115]. While the distinction between single and double sounds "is always potentially significant", the distinction between short sounds and sounds enunciated with true length is not necessarily significant. For instance, the difference between doubled vowels and single long vowels "may often be inappropriate to the ear, and may be even objectively non-existent to listeners, yet the difference would seem to be in most cases subjectively perceptible to the speaker." [p.116].

Jones presents examples from French and German to show that long vowels which are subjectively double may always be pronounced double; whereas with true long vowels a pronunciation with double sounds is not a possible alternative. Therefore, it is not possible to pronounce the French /mɛ:tr/ 'maître' as [mɛ-ɛtr] or the German /ba:n/ 'Bahn' as [ba-an]. He remarks that by reference to the structure of the language, the speaker can decide whether a vowel heard as long is to be considered double. He adds that the distinction between lengthening and doubling is not only difficult to perceive in the case of vowels, it is even still more difficult in the case of consonants. For him the difference between a long consonant and a doubled one is not significant.

From a practical point of view, Jones (op.cit.) suggests that it is feasible to consider "all intervocalic long consonants as double, on the ground that it is usually possible in precise speech to separate them into two by a diminution of force in the middle, attaching the first part to the first syllable and the second part to the second syllable." [p.116]. To put this suggestion into real practice, Jones says that in English, for example, long consonants are divisible when they occur in compound words and in words formed with either prefixes or suffixes, e.g. book-case, pen-knife, oneness, genuineness, solely, etc. He finds it practically preferable to treat long consonants occurring intervocalically, and which are not affected by derivative rules such as that in Arabic,[1] as double consonants rather than as long ones. Initial long consonants are regarded as double "whenever there is reason to consider the first part as a prefix or as an element of a compound." [p.117]. Final long consonants are also regarded as double in all cases "where there is reason to think that the native speaker feels them to be double, and more particularly when the latter part can be regarded as syllabic." [p.118]. Examples are illustrated from languages other than English, since in the latter neither initial long consonants nor final long ones do occur in the way they are here discussed.

Phonetically, in English long final consonants preceded by short vowels, e.g. hill, pen, are never pronounced double. Jones (op.cit.) considers such consonants as merely long. He comes to the final conclusion that the two terms 'long' and 'length' are used only to refer to sounds which have 'true' or 'indivisible' length, and that any lengths which can be regarded as doubling or repetition do not come within the range of this definition.

Pike (1947) seems to be terminologically at variance with the above views. In his reference to the stressed geminate vowel sounds in Hungarian, he does not appear to use the term 'geminate' in his own exposition. Pike uses the term 'double stops' in the same sense as the term 'double articulation' used by other scholars. He states that "For certain sounds, the Double stops, there are two closures in the mouth in addition to the velic closure in the nasal

---

[1] As an example, Jones (op.cit.) gives the word /dikka:n/ 'shop', the plural of which is /daka:ki:n/ 'shops', p.117.

passageway. One of the oral closures may be at the lips and the other at the velum."[p.34]. On the other hand, Pike (op.cit.) makes extensive use of the term 'long consonant', which comprehends 'the long voiceless stop' whose first part tends to serve as an arresting element for the preceding syllable; whereas the second part serves as a releasing element to the following syllable.

Ladefoged (1975), too, presents geminate consonants as a class of long consonants. He describes long consonants (or vowels) which can be analyzed as double consonants (or vowels) as geminates. For him the Italian geminate consonants in words such as nonno 'grandfather' versus nono 'ninth' and folla 'crowd' versus fola 'fable' are the contrasts between English consonants in white tie, why tie and white eye.

Ladefoged (ibid.) remarks that the difference between the geminates in the two languages is that "in Italian a long consonant can occur within a single morpheme ... But in English, geminate consonants can occur only across word boundaries, as in the previous examples, or in a word containing two morphemes, such as unknown or guileless." [p.223]

Trubetzkoy (1969), Martinet (1964) and Cantineau (1960) discuss the concept of gemination from a functional point of view. Trubetzkoy himself uses the term 'geminate (or geminated) consonants' which he believes occurs in many languages. According to him geminate consonants are distinguished from simple or non-geminated consonants by a longer duration, and in most cases also "by a more energetic articulation that is reminiscent of the correlation of intensity." [p.161]. Just like other scholars (e.g. Pike, op.cit.), Trubetzkoy states that geminate consonants, occurring intervocalically, are distributed over two syllables with their first part, i.e. onglide, being grouped with the preceding syllable and their second part, i.e. off-glide, being grouped with the following syllable. He proposes that intervocalic geminates have similar effects to those of consonant clusters on their environment. Trubetzkoy concludes that "All these features point to a polyphonematic interpretation, that is, they call for an interpretation of 'geminated' consonants (or 'geminates') as a combination of two identical consonants." [loc.cit.]

Martinet (op.cit.) and Cantineau (op.cit.) seem to share Trubetzkoy's views on gemination. Martinet, as cited by Rahim (1980), talks about vowel and consonant duration. He notes that the term gemination can be just 'a realization' of simple consonants in certain phonetic contexts. Cantineau, on the other hand, distinguishes between 'long consonants' and 'geminate consonants'. He describes geminate consonants as "those consonants which are characterized by a prolonged articulation and which are generally transcribed by doubling the consonant symbol." (Rahim, ibid., p.200). For him long consonants are only long realizations of simple consonants; while geminate consonants are groups of two identical consonants. Following Trubetzkoy and other scholars, Cantineau (op.cit.) states that when geminate consonants occur intervocalically within a word, the first of the two identical consonants arrests the preceding syllable and the second consonant releases the following syllable.

Kenstowicz and Pyle (1973) define a geminate cluster as "a sequence of identical consonants or vowels." [p.27]. Their study includes a number of different languages, and examples have been quoted from Sierra Miwok (an Amerindian language of California) to consider the alternations between present and past tense stems of three types of verbs where the contrast between single and long consonants exists. Long consonants are represented as sequences of identical segments and not as single segments marked plus long. Two rules have been suggested. The first shows vowel shortening before consonant clusters and long consonants; whereas the second treats long consonants like clusters in the stress rule. Kenstowicz and Pyle believe that long consonants behave like consonant clusters; a fact that "can be explained quite naturally if they are represented as geminate clusters." [p.28]

2.3 Studies on Gemination and Long Versus Short Consonant Differences

In this section we shall discuss previous works that have dealt with the concept of gemination and other related phenomena in languages other than Arabic. Though much attention will be devoted to experimental works, yet some related theoretical, i.e. phonological, works will also be referred to. Most of the relevant theoretical works have been discussed in the previous section of this chapter. The review will be presented, as far as possible, in chronological order.

Sievers, Rousselot and Josselyn

It was mentioned earlier that the question of whether there is any phonetic difference between long consonants and geminate consonants has been debated for some years. Two different views have been prevailing. One view claims that geminate consonants differ from long consonants in that the pronunciation of the former entails two phases. The first phase constitutes a syllable final occurrence of the consonant; while the second phase, which is rather the rearticulated phase, commences the following syllable and thus "constitutes a syllable-initial occurrence of the consonant." (Lehiste et al., 1973, p.131). The other view treats geminate consonants and long consonants as the same and denies the existence of the two phases.

The first view, viz. the two-phase theory, was first introduced by Sievers (1876) who notes that a double consonant is different from a single prolonged consonant. He points out that when implosion-explosion occur within such a double consonant, there is discontinuity of expiration and that the syllable division takes place between the doublets. (Stetson, 1951, p.60). Sievers bases his suggestions on auditory and kinesthetic perception.

Soon afterwards, Siever's claims were subjected to intensive experimental investigations. For instance, Rousselot's investigation (1891), as cited by Lehiste et al. (ibid.), was primarily dependent on extensive use of the kymograph. Using the Gallo-Roman dialect of Cellefroum in his study, Rousselot finds no double articulation of long consonants in that dialect. The results show that the long consonants of that dialect appear about twice as long and twice as intense as the corresponding short consonants. His analysis fails to distinguish between the double, i.e. geminate, and the single consonants. The conclusion he reaches is that when a consonant is prolonged it becomes double. However, neither Rousselot nor Sievers explain the greatly increased length of the double consonant or its influence on the preceding vowel.

At this stage one has to mention Josselyn (1901) who uses Rousselot's methods for studying Italian sounds. As reviewed by Stetson (op.cit.), Josselyn discusses the peculiarities of the Italian double consonants which indicate the increased length of the occlusion and the compensatory shortening of the preceding vowel. His final findings show that the double consonant is greatly different from the single consonant in duration and in function.

Stetson

Referring to previous works, Stetson (1951) maintains that Rousselot's instrumental technique was not sufficiently sensitive to fluctuations at the rate involved nor sensitive enough to bring out the differences in air pressure during the hold of the consonant. He comments that Rousselot's conclusion that "a consonant prolonged becomes double" is entirely inadequate.

Stetson (ibid.) made an extensive study of abutting consonants in English in which intraoral air pressure tracings as well as tongue and lip markers were obtained from a number of tokens grouped in minimal pairs, such as up puppy / lob Bobby;

topic/top pick; hit him/hit Tim; I do/I'd do; un-own/unknown; I lie/I'll lie... etc. While carrying out the experiments, the attention of the subjects was not called to the fact that the same consonant appeared as single and as geminate in each token; instead they were asked to say the word or phrase, i.e. each token, so that the meaning would be clear; besides each subject chose his own rate of utterance.

In general, Stetson (op.cit.) investigates the articulatory differentiation of single and double consonants. He measures the closure durations for single and geminate plosive consonants. The results show that distinct double articulations, with no release between the two consonants, have a median closure duration from onset to release of 200 msec. This type of double articulation normally occurs at utterance rates of two syllables per second. Single consonant articulations are found to have a modal closure duration of 100 to 140 msec and generally occur at syllable rates of 3.5 to 4 syllables per second. The obtained tracings of the examined tokens show that the double consonant is "actually two consonants." [p.61]. This finding is considered as a clear evidence for double articulation. On the other hand, the tracings of the air pressure and the lip and tongue movements exhibit two maxima for the involved abutting consonants. Stetson remarks that "The pressure for the 'double' consonant is a bimaximal curve, showing the arrest of the one syllable pulse and the back stroke of the chest muscles, before the increasing pressure for the release of the second syllable pulse." [loc.cit.].

Palmer and Carnochan

Scholars of the prosodic school (e.g. Palmer, 1957; Carnochan, 1957; Mitchell, 1957)[1] support their investigations with palatograms and kymograms. They tend to consider gemination and non-gemination strictly as phonological terms and not as phonetic, nor as grammatical terms. Mortimer (op.cit.) comments that they also "express valuable phonetic insights gathered in the process of specifying phonological exponents and connecting these with the phonic substance." [p.2]. Palmer (ibid.), for instance, recognizing as all prosodists do that a segmental analysis is often neither the most productive phonologically nor most in accordance with the phonetic facts, sets up gemination for Tigrinya as a prosody of the entire word. He states that the distinction between gemination and cluster is wholly phonological, and is never to be made on purely phonetic grounds. He also adds that "Gemination may, in fact, be treated as a prosody of the entire word. With the abstraction of the prosodic features of gemination, the distinction between single and geminate is redundant, but, for convenience, the distinction of CC and C is retained to symbolize the commutation, in the prosodic system, of gemination and its counterpart 'singleness of consonant' or 'non-gemination'." [p.140]. He believes that the prosody of the entire word can have a multiplicity of exponents, of which consonantal duration is only one, occurring singly or in various combinations.

In his investigation, Palmer (ibid.) reproduces

---

[1] Mitchell's investigation has been discussed in detail in section 1.6.2.

palatograms to exemplify the exponents of gemination in Tigrinya. His palatograms illustrate various interdental and alveolar types relating to the exponents of gemination he sets up. Both the vowel qualities of the second syllables and the features of the following consonants are considered as co-exponents of gemination. And since the main interest was solely that of gemination, Palmer was involved in the recognition of the two kinds of plural form found in the language; namely 'suffixed plurals' and 'broken plurals'. For example, for plural forms of the structure CVCvCVC/CVCvCCVC, in which the exponent of the term of the V system of the final syllable is half close central, the exponents of gemination might be stated in terms of:

(i) consonantal duration;

(ii) tenseness of articulation (plosion as opposed to friction);

(iii) position of articulation.

Hence, Palmer (op.cit.) remarks that "It is, perhaps, necessary to point out that where the exponent of gemination is greater duration, the exponents of the phonematic units are to be stated as e.g. 'voiceless labio-dental fricative', but where the exponent of gemination is tenseness, the description is e.g. 'voiced bilabial' since the features of plosion or friction are exponents of the prosody, and where the exponent of gemination is given in terms of position of articulation, the exponent of the phonematic unit is stated as e.g. 'voiceless dental/interdental plosive', since the precise position of articulation is an exponent of the prosody (in this case dentality is the exponent of gemination and interdentality of non-gemination)." [p.141]. At the same time, for plurals of the structure CVCvCV/CVCvCCV paired with the singular CVCCV the exponent of gemination is again to be stated as greater duration only.

Carnochan (op.cit.), on the other hand, studies the phenomenon of gemination in Hausa. He intends to emphasize the relation between observable phenomena and phonological statement. He does not aim to present general theory or to establish universal categories to be regarded as gemination and non-gemination criteria. Although Carnochan presents data provided by instrumental techniques; viz. the kymograph, he believes that "phonological categories cannot be set up purely by reference to instrumental data." [p.149]. This was his justification behind including material from other languages, besides Hausa, such as from English, Yoruba and Twi. The analysis of these languages shows that the velar articulations of English in Leaking, and of Yoruba in Pako 'a plank', and of Twi in Sika 'money' are associated with a single consonant unit in the phonological structure. This, he believes, makes it difficult "to consider seriously the possibility of establishing gemination as a universal category on the criterion of duration, either by listening, or by instrumental means." [p.150]. He also points out that gemination and non-gemination cannot be universally defined, and cannot be said to 'exist' or to 'occur'. The kymograms of three English sentences have been presented. The three sentences are:

(a) I was amazed at the price.

(b) Only for a slim man.

(c) I was amazed!

According to Carnochan (op.cit.) the above three examples "emphasize that duration is not a satisfactory criterion for establishing the phonological category of gemination within a given language." [p.150]. He states that any analysis of English would associate gemination with the bilabial nasal in example (a) as impossible, but not impossible with regard to that mentioned in example (b). The analysis is even more difficult for the bilabial nasal in example (c) whose duration is even longer than that of example (b). He adds that the examples he quoted from English, Yoruba and Twi would suffice "to justify the great need for distinguishing carefully between phonetic phenomena involving longer or shorter durations of consonantal articulation on the one hand, and the establishing of the phonological categories of gemination and non-gemination on the other." [p.154].

Carnochan's discussion of gemination in Hausa is totally based on the examination of examples representing verbal pieces and nominal pieces in the language. The final conclusion of the examined structure is that gemination in Hausa should be considered as a term in 'system' as well as an element in 'structure'. It also shows what correlations are established between the elements of structure and the phonic data of the utterances.

Hegedüs

Hegedüs (1959) studies single and geminate consonants in Hungarian. He examines intensity curves reproduced from tape-recorded utterances which represent items containing single and geminate consonants occurring in word - interior intervocalic positions rather than at word boundaries. His findings give no evidence in the intensity curves to support the two-phase theory suggested by Sievers (op.cit.) and reiterated by Stetson (op.cit.). This is to say, Hegedüs finds no evidence for the existence of re-articulation in Hungarian geminates.

Pickett and Decker

An interesting listening experiment was carried out by Pickett and Decker (1960) to examine the perceptual distinction between the single voiceless plosive [p] and its geminate counterpart [p-p] as a function of closure duration as well as rate of utterance of the surrounding sentence. The two experimenters recorded on tape sentences containing single plosives which could then be artificially lengthened so as to make it sound double and the sentences to differ in meaning. The two sentences He was the topic of the year, and He was the topic of the discussion were chosen for the listening tests. They were recorded on tape and then, in a number of recorded copies, the duration of the /p/- closure was artificially altered by inserting or removing tape. Each of the recorded sentences was copied ten times and then the /p/- closure of each copy was lengthened or shortened to provide a set of ten sentences with ten different closure durations.

The closure was located by listening to the tape output as the tape was pulled slowly by hand across a playback head. Shortening was made by removing tape from the middle of the closure, and lengthening was made by inserting blank tape at the centre of the closure. Each of the altered sentences was judged by a group of listeners to be either He was the topic of the year or He was the top pick of the year, in the case of the first test sentence, or to be either He was the topic of the discussion or He was the top pick of the discussion, in the case of the second one. The effects of ten closure durations were examined in various combinations with five different rates of utterance spanning a range from very slow to very rapid utterance. The rate was paced with the aid of the sweep hand of a 6-rpm stopwatch. The subject watched the moving hand and spoke the sentence at a rate to occupy about the desired times which were calculated on the basis of eight syllables for the sentence. Rates of 2, 3, 4, 6 and 8 syllables per second were chosen for testing, and new recordings of each subject were then prepared by splicing. A threshold closure duration was defined to be the duration at which 60% of the judgments were topic. As the rate increased from 2 to 8 syllables per second, the threshold closure duration decreased from 320 to 140 msec and at a progressively declining rate. This function of threshold closure duration versus rate of utterance was found to be approximately parallel to the relation, for the unaltered sentences, between actual closure duration and rate of utterance.

The test sentences were recorded by the two experimenters who "practised and recorded repeatedly so as to provide final examples having three characteristics: (1) a relatively standard rate of utterance, (2) a /p/- closure of 150 msec, and (3) a relatively even syllable stress throughout the sentence." [p.12]. The results of the listening tests with normal rate of utterance show that generally most /p/- closures shorter than 150 msec were judged as single consonants and closures longer than 250 msec were, on the other hand, judged as double consonants. The results also show an interaction between rate of utterance, the perception of the consonant, and its duration. It was found that as rate became slower, a larger duration was necessary for the consonant to be heard as double; whereas with short durations and fast rates, the interaction seemed less marked. Picket and Decker (op.cit.) contend that these results are quite consistent with the articulatory measures reported by Stetson (op.cit.). The tests also show that there is a bias towards making single consonant judgments, even when closure durations are at their greatest; besides the two talkers give "very similar relations between closure duration and single consonant judgments." [p.13]. They attribute this to their failure in attaining a style of utterance having no bias towards the single consonant perception. Mortimer (op.cit.) comments on this saying that "a possible additional contributory factor not mentioned by the experimenters, is that as <u>topic</u> seems to be a more frequent and likely item than <u>top pick</u>, listeners may be biased towards the former, even in a forced choice situation.

And it may, anyway, be true that a general bias towards the selection of non-geminate items exists." [pp.6-7].

At the same time, the results of the tests with various rates of utterance show that there is an interaction between rate of utterance, the perception of the consonant, and its duration. Evidence is found that as rate becomes slower, a longer duration is necessary for the consonant to be heard as double. However, with shorter durations and faster rates, the interaction seems less noticeable. In this respect, Huggins (1972) points out the basic limitations in Pickett and Decker's experiment. He comments that the only alternative responses available to the subjects are either topic or top pick. The former contains an unaspirated medial [p]. All stimuli are derived from a production of the word topic; therefore none of the stimuli are aspirated. Huggins concludes that the discrepency between his experiment and that of Pickett and Decker is due to the fact that in the latter experiment "the cue of closure duration had to override the conflicting cue of no-aspiration before subjects would perceive the stimulus as top pick ... " [p.1277].

Mieko Han

The feature of duration plays an important role in Japanese phonology. It is well known that in Japanese a long vowel contrasts with its short cognate, and similarly, a long consonant contrasts with its short partner. Consonant gemination is an interesting phenomenon in Japanese, especially when it occurs in European loanwords. For instance, Han (1965) studies the feature of duration in Japanese. The purpose of his study is to present the results of experiments made by means of a spectrograph on Japanese vowels and consonants, and to determine their inherent duration. The five Japanese vowels [a i u e o] are pronounced in the same phonetic environment, and syllables such as /ka/, /ki/, /ku/, /ke/ and /ko/ are spoken with constant tempo. In order to measure the duration of Japanese vowels in the same phonetic context, these syllables are enunciated with the addition of an extra syllable inserted at the beginning and at the end of each utterance. This makes it possible to pronounce the five vowels free from intonational influence. For example, the sequence of syllables /kakikukeko/ is pronounced as /kakikukekokane/, and /naninuneno/ as /naninunenonano/. Thus all five vowels are pronounced in as similar phonetic environment as possible.

A second experiment was also conducted in which the same five syllables were placed in a different order so as to make a meaningful utterance. Sentences such as /kakikakeno o kakukotoda/, where the five Japanese vowels occur in a similar context on the same pitch, were constructed. Han (op.cit.) observed that vowel [u] was always the shortest among the five vowels if it was pronounced in the same environment; [i] was the next shortest and [o], [e] and [a] followed in that order. The duration ratio among the five vowels was reported as in table (2.1) below:

Table (2.1)

Vowel	Ratio value
[u]	1.00
[i]	1.17
[o]	1.26
[e]	1.37
[a]	1.44

(After Han, 1965, p.67)

Similarly, the duration of Japanese consonants was measured on the spectrograms. The results show that in the same phonetic context voiceless consonants are relatively longer than their voiced counterparts. For example, spectrographic inspection indicates that there is approximately 20 to 40% increase in duration in the voiceless plosive [p] in /pan/ 'bread', [t] in /hitai/ 'forehead', [k] in /seekaku/ 'personality', [tʃ] in /hatʃida/ 'it's a bowl', and [s] in /susume/ 'march forward', over their corresponding voiced partners [b] in /ban/ 'turn', [d] in /hidai/ 'over weight', [g] in /seegaku/ 'vocal music', [dʒ] in /sandʒaku/ 'three jakus', and [z] in /suzume/ 'sparrow'.

Han (op.cit.) groups the Japanese consonants into five classes in accordance with their relative duration, ranging from the longest to the shortest. Voiceless fricatives and plosives are considered as the longest; while the approximants are regarded as the shortest. Other classes come in between. He remarks that there are minor differences in duration among different consonants belonging to the same class, and in the event of geminate plosives, i.e. [pp], [tt] and [kk], the minor difference, due to the place of articulation, is clearly observable. The durations of short and long consonants in words such as /supai/ 'spy' versus /suppai/ 'sour', /ita/ 'existed' versus /itta/ 'went', /haka/ 'temb' versus /hakka/ 'mint', /ise/ 'a place name' versus /isse/ 'a unit of area', etc., are measured and it is found that the duration of short and long consonants is, on the average, in the ratio of 1.0 to 2.6 and often 1.0 to 3.0. This indicates that long consonants are genuinely much longer than twice the short cognates in Japanese.

According to Han (op.cit.) the length of a Japanese vowel or consonant segment is not the basic unit of duration. He claims that his spectrographic experiments correct the notion of double consonants or double vowels in Japanese. His analysis of a number of spectrograms convinces him that there is, in Japanese, a unit of duration associated with the syllable which he calls "onsetsu". The term onsetsu is "functionally a quantum of duration." [p.71], and Japanese utterances are made up of one or more onsetsu. It is taken to mean a sound-unit in Japanese. Han points out that because the actual length of each onsetsu is approximately the same, Japanese is characterized by its staccato quality.[1] There-

---

[1] In his paper on Japanese phonetics, Block (1950) states that "The most striking general feature of Japanese pronunciation is its staccato rhythm. The auditory impression of any phrase is of a rapid patterning succession of more-or-less sharply defined fractions, all of about the same length." [pp.90-91]. It is of interest to comment here that some aspects of Japanese prosodies give the impression to speakers of other languages of having a staccato quality. The term 'staccato' can be used in the description of a language's phonological system relative to some speakers' norm. (Crystal, 1969, pp.161-167). But in the above sense, it is used in a less rigorous way to refer to the speaker's impression of other languages.

fore, he proposes a new phonemic entity of onsetsu duration which he symbolizes by a /:/, and "may be added to a consonant that has one onsetsu duration by itself." [p.72]. Consequently, long consonants normally transcribed as [pp], [tt], [kk], etc., would be phonemically represented as /p:p/, /t:t/, /k:k/, etc., respectively.

Delattre

Delattre (1971) made probably the most stimulating and the most wide-ranging study of the phenomenon of gemination. His investigation includes material from four languages, namely English, German, French and Spanish. He differentiates between the two terms 'double' and 'geminate'. The former "is usually reserved for graphic symbolization of two consonants." [p.31]. For him the concept of gemination applies to the meaningful perceptual doubling of a consonant phoneme.

Delattre (ibid.) bases his study on spectrographic and cineradiographic examination of utterances in the four languages. The study is complemented by using speech synthesizers in order to "test by ear what consonant durations, or ratios of consonant durations are appropriate in each language in distinguishing geminate consonants from single ones." [p.32]. The objective of this investigation is to examine "the acoustic and the articulatory correlates of consonant gemination, both across and within word boundary, and to compare their behaviour among four languages." [loc.cit.]. The selected stimulus items are appropriate utterances composed in each language in the form of minimal pairs or near minimal pairs so that the variables may occur in similar phonetic environment. They are chosen so as to give consonant gemination either across word boundary, as in English <u>We lend</u> or <u>We'll end</u> versus <u>We'll lend</u>, or within word boundary, as in Spanish <u>Caro</u> versus <u>Carro</u>. The utterances are recorded by native speakers of each language, and spectrograms are made to measure consonant duration and observe the tempo of formant transitions and the variations of overall amplitude.

Spectrographic inspections indicate that consonant duration is an important factor in the linguistic functioning of gemination. The final results reveal that consonant duration is a major attribute of gemination across word boundary in all four languages, but that the duration contrasts are wider in the two Latin languages than in the two Germanic ones, and narrowest of all in English. This is interpreted as indicating that "consonant gemination across word boundary is less distinct, less stressed, more slurred in English or German (especially in English) than in French or Spanish." [p.38]

Delattre (op.cit.) points out that there were cases in which one of his subjects made very little distinction in duration between a geminate and a single consonant; still, auditorily, the longer consonants were clearly heard as geminates. On the other hand, there were cases in which very long single consonants were clearly heard as merely single. This, perhaps, justifies Delattre's assumption that there are other factors, in addition to duration, that contribute to the production and perception of gemination. Therefore, intensity variations, as for example in the case of [n], [l] and [s], are found to play a definite role in distinguishing geminates from single consonants. They exhibit two phases in the articulation of geminate consonants for all four languages; one with the features of final consonants, and the other with the features of initial consonants. The cineradiographic evidence indicates that the first phase is marked by consonant anticipation and weak tongue pressure; whereas the second phase is marked by vowel anticipation and increased tongue pressure. The perceptual tests, carried out with synthetic speech, confirm that consonant duration is a major cue for the perception of gemination, and suggest that German, French and Spanish ears require a longer hold than American ears in identifying a consonant as being geminate, across word boundary. (Delattre, op.cit., p.113). The tests also denote that the intensity of arresting and releasing formant-transitions is a major cue for the perception of gemination, and one likely to be overridden by the duration cue.

Lehiste, Morton and Tatham

Earlier, it was mentioned that Hegedüs (op.cit.) studied intervocalic geminates in word-internal positions and found no evidence of double articulation, and that Stetson (op.cit.) and Delattre (op.cit.) examined consonants abutting at word boundary and, conversely, found evidence of double articulation. Lehiste et al. (1973) made a further attempt and studied both situations within the framework of a single investigation. The aim of this investigation was "to establish whether the position within a word or at word boundary has any influence on the production of geminate consonants." [p.132]. The languages chosen for the study were English and Estonian, and the technique used was electromyography, supplemented by intraoral air pressure measurements, because it was assumed that this technique could provide an unexplored opportunity for finding out whether gemination is connected with rearticulation and that the electromyographic signal "would indicate the duration and extent of the electro-chemical activity within the muscle consequent upon innervation of that muscle. The intraoral air pressure was sampled to detect accurately the moments of closure and opening of the lips - i.e. the stop phase of the bilabials [p] and [b]." [p.134]. Two subjects, one a native speaker of Estonian and the other a native speaker of American English, acted as informants and each pronounced both the Estonian and English sets of materials used in the study. Besides investigating the difference between single and geminate consonants in word-internal position and at word boundary, the study also aimed at examining the difference between short and long geminates, the difference between the manifestations of geminates at word boundary in the two languages, and the differences between productions of geminates by native and non-native speakers of the two languages. This is probably the reason behind choosing

Estonian for this investigation, since it is a particularly suitable language for studying geminates and has not only an opposition between short and long consonants in intervocalic position, but also an opposition between short and long geminates. The examined consonants are voiced and voiceless bilabial plosives, presented in sets such as taba, tapa, tappa, lapp peal, for Estonian, and lee pin, leap in, leep pin, for English. The Estonian consonants are classified as follows:

(a) short consonant = single
(b) long consonant = short geminate
(c) overlong consonant = long geminate

Lehiste et al. (op.cit.) point out that closure duration serves effectively in distinguishing among the three intervocalic consonant lengths in Estonian, i.e. short, long and overlong (as spoken by a native speaker), and in differentiating between pre- and post-junctural occurrences of consonants in English (as spoken by a native speaker). They add that it plays no part in distinguishing geminate from single consonants in English, and that its role is inconsistent in distinguishing between medial overlong consonants, i.e. long geminates, and junctural geminates in Estonian. It appears that electromyographic amplitude plays no practical role in distinguishing geminates from other occurrences of the consonants included in the investigation. For instance, initial consonants in English are found to have less electromyographic peak amplitude than final ones, and differences in peak electromyographic amplitude in the Estonian data. This is to say, consonants included in the investigation cannot be distinguished from each other instrumentally on the basis of electromyographic amplitude owing to the fact that electromyography could not possibly be a factor in perception of consonantal differences.

The ultimate conclusion reached by the investigators is that "articulation of geminates is a language specific phenomenon not easily duplicated even by a skilled phonetician." [p.147]. Differences between short and long medial geminates in Estonian word-interior positions as well as at word boundaries are firmly established. Evidence, based on differences in the number of electromyographic peaks, is found for rearticulation in the production of Estonian intervocalic long and overlong consonants as well as in consonantal sequences at word boundaries. On the other hand, there is no evidence for rearticulation in the event of English consonantal sequences at word boundaries. Lehiste et al. (op.cit.) reported that "electromyographic duration and number of electromyographic peaks distinguished Estonian short and long geminates and English pre- and post-junctural voiceless plosives. Significantly, number of electromyographic peaks did not differentiate between single consonants and junctural geminates in English." [p.147].

<u>Sampson</u>

Sampson (1973) studied consonant duration in biblical Hebrew. He disagrees with other researchers (e.g. Lehiste, 1970) who claim that duration contrasts should be represented with geminate versus single segments. Instead, he argues that duration contrasts in biblical Hebrew consonants are to be analyzed as long versus short rather than as clusters of identical consonants, i.e. geminates versus single consonants. His argument is based primarily on the difficulty one would have in formulating a rule which spirantizes obstruents post-vocalically, and therefore possibly pre-consonantly, if geminate clusters were posited. According to him, spirantization does not apply where there is an identical sequence of consonants.

Sampson (op.cit.) argues for interpreting long consonants as single [+ long] segments on the basis of the fact that spirantization in Hebrew must be greatly complicated if long consonants are instead treated as geminates. This rule, then, converts post-vocalic plosives to fricatives, even if the plosive in question is the first member of a consonant cluster. Sampson's arguments, however, were later strongly criticized by Barkai (1974) who commented that Sampson's rules do not apply in several positions.

Klatt

Klatt (1974) measured the duration of the phonetic segment [s] in English words. A list of more than one hundred words was constructed in such a way that this segment appeared in several different phonetic contexts. In the majority of these words, the segment [s] appeared as a single consonant in three different word positions in words containing different numbers of syllables and different stress patterns. Words containing the double [s] were also included and the duration differences between single and double [s] were compared.

The final results indicate that the duration of [s] depends on whether the following vowel is stressed or not. An [s] is found to be longer in prestressed position and shorter before unstressed vowels in word-final position, and that it is shorter in multisyllabic words and in consonant cluster sequences. Klatt (op.cit.) states that "the relatively large durational variability of [s] before an unstressed vowel suggests that unstressed syllables are articulated with less precision than are stressed syllables and supports the notion that lexical stress is a feature the domain of which encompasses both consonants and vowel." [p.51]. The results also indicate that the double [s] is longer than a single [s] for all subjects in English words such as misspell versus misplace and misstep versus mistake. A double [s] followed by a plosive is found to be approximately 25% longer than a single [s] followed by a plosive.

Fujisaki, Nakamura and Imoto

An experimental investigation was performed by Fujisaki et al. (1975) both on subjects with normal hearing and on hard-of-hearing children. The investigation primarily aims to examine "the roles played by segmental durations in the perception of Japanese vowels and consonants in various contexts, and their relationships to the perception of duration of various non-speech sounds which possess acoustic features similar to the speech sounds under study." [p.198]. Unlike Han (op.cit.), the investigators use the term 'geminate' in place of the term 'long' for consonants.

Japanese voiceless consonants, used intervocalically, contrast in duration with their longer counterparts and are often denoted by the juxtaposition of two identical phonetic symbols. Fujisaki et al. (op.cit.) state that in voiceless fricatives, it is the duration of the quasi-stationary friction that constitutes the contrast; while in the case of voiceless plosives and affricates, on the other hand, the phonemic contrast is realized by a difference in duration of the stop gap preceding the plosion. Word-medial nasal consonants may be changed to their longer cognates by "an increase in duration of the nasal murmur." [p.199]. They observe that the duration of a long vowel or a geminate consonant is more than twice as large as its shorter counterpart.

The stimulus items of their first experiment were selected in such a way that "they should have the same type of pitch accent, and should contain the segment under study in the minimal context, so that they could be discriminated solely by the durational cue." [p.207]. The test results of these subjects were analyzed individually. They show marked uniformity of phoneme boundaries and accuracies of identification for vowels, nasals and voiceless plosives regardless of the spectral features of segments and their contexts. In order to obtain quantitative estimates for the magnitude and the extent of influences caused by the segmental durations or the talking rates of their context, further identification tests were performed using synthetic speech stimuli of the fricative group both in word and in sentence context, but at different tempos.

Similar tests were conducted on four hard-of-hearing children aged between 11 and 12 years with severe sensory-neural hearing impairments. Three other children with normal hearing, of the same age, were also tested as the control. Summarizing, the results of these tests denote that despite their severe hearing impairments, none of the four hard-of-hearing children differ significantly from the three normal ones in their perceptual ability of duration of non-speech stimuli when their hearing loss is compensated by proper amplification. However, their performances in perception of speech stimuli exhibit extensive individual variations, ranging from no significant difference from the normal group at one extreme to highly significant differences in all the tests at the other. (Fujisaki et al., op.cit., p.227). The investigators conclude that these individual differences are found to be dependent more heavily on previous history of education than on types and degrees of hearing loss, and testifying to the importance of proper training at earlier ages.

Thananjayarajasingham

Thananjayarajasingham (1976a) used direct palatography for his study of intervocalic geminate consonants in Ceylon Tamil. The examined items consist of words in which single and geminate nasals and laterals occur. They are all nominal forms which frequently exist in the colloquial Tamil speech of the experimenter who is a native speaker of Ceylon Tamil and who acted as his own subject. The obtained palatograms were divided into various zones in accordance with the Firthian lines. (Firth, 1948). They reveal a clearer and more marked wipe-off in the case of words containing geminate consonants than in those words containing their non-geminate partners. This, of course, indicates that during the articulation of geminate nasals and laterals in Ceylon Tamil, as spoken by the experimenter, the tongue touches the area of contact more firmly and for a longer duration than in the articulation of single nasals and laterals.

In a subsequent study, Thananjayarajasingham (1976b) used kymography and spectrography as two experimental techniques for measuring the duration of intervocalic voiceless geminate plosives occurring in colloquial Ceylon Tamil, as in the words tappu 'mistake', pattu 'tin', kaattu 'wind' pookku 'behaviour', ... etc. The kymographic and spectrographic evidences obtained from this experiment indicate that "the duration of the closure was very long if the stop was preceded by a short vowel and considerably less if the stop was preceded by a long vowel." [p.150].

Mortimer

In 1977, Mortimer made a phonetic study to investigate certain acoustic and perceptual aspects of grammatical distinctions in English from a pedogogical point of view. The investigator had noticed that students of English as a foreign language experienced difficulty in hearing auxiliary verbs when these were signalled by a single consonantal sound. The difficulty was greatest if this sound was homorganic or identical with that at the beginning of the verb root. Therefore, the study aimed at finding out how far native listeners would be able to distinguish past simple from present perfect, past simple from past perfect, and present simple from future simple when these tense distinctions were minimally signalled without additional grammatical or semantic cues. More specifically, the investigation was concerned "with the extent to which native listeners perceive particular manifestations of grammatical tense distinctions under various conditions." [p.18]. Paired utterances, such as He sold it/ He's sold it, He dropped it/He'd dropped it, and He let it/ He'll let it, were chosen and were embedded in three constructed monologues which were recorded by the investigator in either a dramatic or narrative style. Ten undergraduates, who were all native speakers of English, participated in this study as speaking subjects. An experienced phonetician agreed to undergo the full elicitation, following exactly the same procedures used with the ten undergraduates. The ten undergraduates as well as the phonetician were completely unaware of the nature of the project.

Mortimer (op.cit.) was interested in finding out whether there were measurable, consistent durational differences in the acoustic structure correlatable with the geminate/ non-geminate distinction at the critical point of juncture, and whether measurements at the point of juncture correlated with measurements of total utterance duration. He was also interested in finding out whether the acoustic pattern at the point of juncture was readily divisible into two segments correlatable with post and prevocalic consonants respectively, and whether variability in the duration and/or formant structure of the initial [i] segment of the experimental phrases was correlatable with the non-geminate/geminate contrast.

Scrutiny of spectrograms indicates that all geminate members of a pair have a greater acoustic segment duration than do their non-geminate partners in that pair. The geminates [z̥s] and [ll], for instance, tend to be about a third longer than the non-geminates [s] and [l], but [dd] is found to be usually longer by about 75% than [d]. On the other hand, there is little evidence in the data for regarding [z̥s] and [dd] acoustically as two segments; whereas there is ample evidence for an internal segmentation of [ll]. The spectrographic inspection also shows that the vowel [i] tends to be longer in pre-geminate positions than in pre-non-geminate positions in the [d]/[dd] pair. A similar but stronger tendency is still seen in the case of the [s]/[z̥s] pair; whereas the [l]/[ll] pair is not "amenable to investigation in this regard." [p.54].

The listening tests, on the other hand, indicate that non-geminate items are better discriminated than geminate ones. However, when non-geminates are presented in a forced choice, they are extremely well discriminated by almost all listeners. The geminate/non-geminate distinction is, on the whole, well perceived when listeners are aware that this distinction is at issue despite the fact that perception varies according to speaker and according to manner of articulation of the consonants at the point of juncture. Mortimer (op.cit.)

adds that "on the basis of the limited exploratory experiments conducted in which subjects were unaware of the issue, however, it is tentatively concluded that though listeners have the capacity to discriminate most geminates, a geminate tends to be perceived as a non-geminate unless the means of discrimination are somehow altered by an aspect or aspects of context." [p.116].

Table (2.2) Average durations (in msecs) of non-geminate/ geminate segments for all speakers over all runs.

Non-geminate		Geminate	
segment	duration (in msec)	segment	duration (in msec)
[s]	120	[z̥s]	145
[d]	60	[dd]	115
[l]	125	[ll]	180

(After Mortimer, 1977, p.44)

Table (2.3) Average durations (in msec) of vowel [i] preceding non-geminate versus geminate consonants.

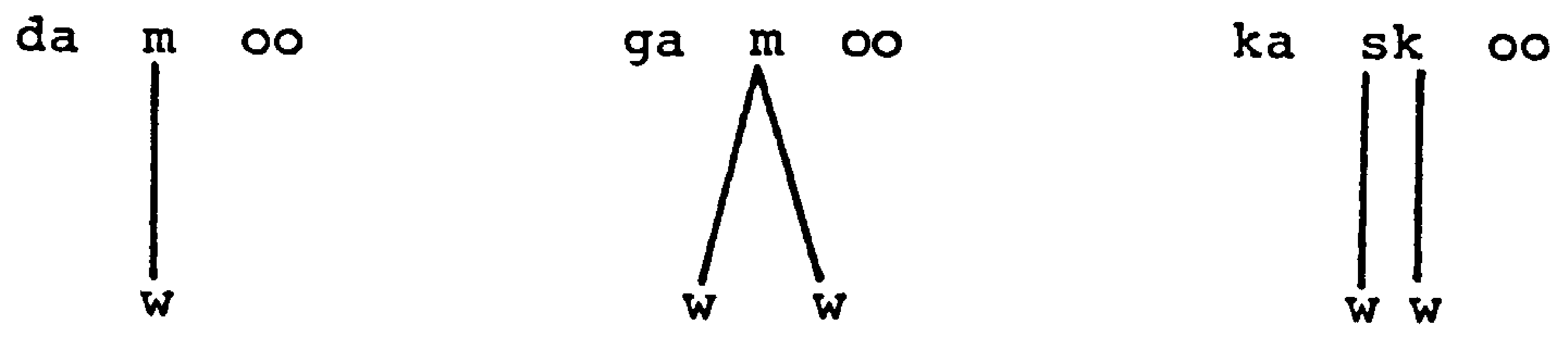
Segment	[i] before non-geminate	[i] before geminate
[s]/[z̥s]	10	36
[d]/[dd]	13	23

(After Mortimer, 1977, p.53)

Leben

Following Sampson (1973), Leben (1980) subjected biblical Hebrew to further investigation in which the problem of segmental quantity cast in terms of a metrical theory of syllabification was studied. Justification for this theory comes from consonant quantity paradoxes in Hausa and biblical Hebrew, and instances of long consonants as single segments and as clusters are examined. Leben (ibid.) terms the level at which quantity is expressed as the 'metrical level', and proposes that the units of quantity are *s* for 'strong' and *w* for 'weak'. For him, *s* also corresponds to [+ syllabic] and *w* to [- syllabic]. A single unit of length, *w*, is associated with a short consonant; while two units of length, *ww*, are associated with a long one. Besides, a sequence of short consonants is associated with a sequence of *ws*.

The sketch below illustrates how Leben (ibid.) equates long consonants with short ones at the segmental level while identifying long consonants quantitatively with consonant sequences at the metrical level. Examples are from Hausa:[1]


a) short C b) long C c) CC cluster

Reviewing previous work (e.g. Sampson, op.cit.; Barkai, op.cit.) Leben finally states that the metrical approach provides a means for formalizing the observation

---

1. See Leben, 1980, p.499, for further details.

reported by Barkai that in biblical Hebrew "generally a vowel is long before a single consonant and short before a consonant cluster including clusters of identical consonants." [p.458]. He adds that a vowel preceding a cluster or long consonant will be a syllable of the pattern (CVC), with the metrical values wsw; whereas a vowel preceding a single short consonant followed by a vowel will be in a syllable of the pattern (CV), with the metrical values ws.

McKay

McKay (1980) carried out a spectrographic investigation of word-medial plosive geminates in Rembarrnga, a north Australian language, to determine the legitimacy of interpreting medial plosive contrasts in this language as contrasts between geminate and single plosives, rather than between voiceless and voiced plosives. He tried to ascertain whether the acoustic facts were compatible with the gemination interpretation. Phonologically, voiced and voiceless plosives in Rembarrnga are in partial 'complementary distribution', contrasting only in medial position, except medially after nasals. Morphophonemic analysis shows that there is "fairly strong evidence from the use of various affixes in support of the gemination interpretation. Where affix and stem bring together two identical stops the result is a long, voicelss, fortis stop. If only one (syllable initial) stop occurs it is short, voiced and lenis." [p.344].

The final results drawn from this experiment indicate that the durations of occlusion for the geminate and single plosives show clearly that the geminates are considerably longer than the singles. The differences in closure duration between geminate and single velar and bilabial plosives found in this study are comparable with those reported by Lehiste et al. (op.cit.) between single plosives and short geminate plosives in Estonian. In some cases, vowels are found to be insignificantly shorter before the geminate plosives than before the single ones. McKay (op.cit.) draws the conclusion that "structural and spectrographic evidence appears to indicate that the opposition between medial stops in Rembarrnga - short lenis versus long fortis - can and should be interpreted as an opposition between single and geminate stops." [p.346].

Table (2.4) Mean durations of medial plosives (in msec) in Rembarrnga. (n) represents number of investigated tokens.

	p	t	ṭ	ƫ	k
Geminate					
Mean	130.6	193.3	125.4	230	127.2
S.D.	21.44	17	40.98	29.44	26.65
n	15	3	11	3	11
Single					
Mean	62	35	21	66	60
S.D.	14.69	5	5.39	8	28.28
n	15	4	10	5	10

(After McKay, 1980, p.345).

Paul Ken Aoki

Aoki (1981) made a comprehensive investigation whose aim was to determine the nature of phonological processes creating geminate consonants in native Japanese words and the extent to which these processes are applicable to European loanwords entering Japanese. For example, when the English word hit is borrowed in Japanese, it will appear as /hitto/.

However, in examining native Japanese phonological alternations, one notices that geminate consonants do occur. Aoki (ibid.) does not agree with previous analyses describing such consonants as the consequence of two "morphologically juxtaposed consonants existing in underlying representation and undergoing assimilation processes, /kir+ta/ (verb stem + past tense)——►[kitta] 'cut'." [p.1]. He is, in fact, in disagreement with those researchers who claim that geminate consonants are divided at morpheme boundaries and that they surface by means of consonant assimilation. He remarks that "If geminate consonants are always divided by morpheme boundaries and are a result of consonant assimilation, it is not apparent why geminates should appear in loanwords, especially in light of examples like [hitto] - 'hit' where it is clear that there is neither an internal morpheme boundary nor an underlying source consonant such that consonant assimilation can apply, yielding a geminate [tt], e.g. /hiCt/." [loc.cit.].

Aoki (ibid.) states that, in general, Japanese geminate consonants occur only intervocalically and appear to have the same pattern as that of the other consonant sequences, namely a syllabic nasal followed by a consonant, which also only occurs intervocalically. Geminate consonants existing in native Japanese comprise these combinations: [pp, tt, kk, ss, ʃʃ, tʃtʃ]. It is quite evident, then, that Japanese has no voiced geminates; they are all voiceless. The first part of a geminate consonant terminates the first syllable and the second part initiates the following one. This is why loanwords, such as those of English origin, containing voiced geminates are not considered easy for native Japanese speakers to pronounce. Words, such as /dabbu/ 'dab', /beddo/ 'bed' and /doggu/ 'dog', are very often pronounced as /dappu/, /betto/ and /dokku/ respectively, where a voiced geminate is replaced by its voiceless equivalent. On the other hand, words like /koppu/ 'cop' and /kappu/ 'cap' are easily pronounced since they originally have voiceless consonants. In this matter, Aoki employs Jaeger's (1978) hypothesis claiming that voiced geminates are physiologically difficult to articulate and, therefore, they should either be prenasalized or imploded.[1] In the case of Japanese, he suggests that the first part of a geminate becomes nasalized. He considers this suggestion as the only hypothesis that "accounts for the apparent absence of voiced geminates in native Japanese". [p.171]. Aoki also maintains Hyman's (1975) proposal that loanwords are lexicalized in terms of the underlying forms of the target language which then undergo the rules

---

1. Jaeger (1978) states that "Voiced geminate obstruents are more unstable than voiceless; over time they should either devoice, or devise some means of prolonging voicing, for example, becoming prenasalised or imploded." [p.320]. In as far as the British terminology is concerned, the word 'imploded' would then stand for 'implosive'.

of the phonology of that language. According to Hyman, foreign words and segments are perceived in terms of the phonology of the native langauge, as in these monosyllabic words which are borrowed by the Japanese with geminates:

/hitto/	'hit'
/kappu/	'cap'
/bukku/	'book'

Balasubramanian

Closely following Thananjayarajasingham's (op.cit.) two experiments for examining the doubling of intervocalic nasal and lateral consonants in Ceylon Tamil, Balasubramanian (1982) started an investigation to examine the same feature in colloquial Indian Tamil. Direct palatography and electrokymograph were exploited in this study to provide phonetic evidence. The test items were presented in minimal and/or sub-minimal pairs indicating contrast between single and double nasal and lateral consonants. These items were "carefully chosen to ensure that only one segment in each word was capable of wiping off the marking medium from the palate." [p.99]. Four informants, who were native speakers of Indian Tamil, participated in this study. All palatograms obtained were divided into various zones following the Firthian lines and similar to those produced by Thananjayarajasingham (op.cit.).

Balasubramanian (ibid.) points out that the palatograms of words containing single nasals or laterals show a mere suggestion of a wipe-off; while those of words having their double cognates reveal a clearer and wider wipe-off.

These evidences clearly illustrate that the tongue makes a firmer contact with the palate for a longer duration during the articulation of the double consonants than during the articulation of their single counterparts in Indian Tamil. Besides, the kymographic tracings show that the duration of the double nasals and laterals is about two and a half times that of their single partners. The findings of this investigation are in close agreement with those findings reported by Thananjayarajasingham (op.cit.) regarding the same phenomenon in both Tamil dialects.

Table (2.5) Average duration (in msec) of 160 utterances contrasting between single and geminate Tamil consonants used intervocalically.

consonant checked	single between vowels	geminate between vowels
[m]	125	285
[n]	122	290
[l]	140	360
[ɭ]	145	380

(After Balasubramanian, 1982, p.103)

2.4 The Relationship Between Word-Medial Consonants and the Duration of a Preceding Vowel:

The relation between consonants and the duration of the preceding vowels has been subjected to thorough investigations for a considerable length of time. Researchers seem at variance whether the duration of vowels is essentially influenced by the type of the following consonants. In this section we shall discuss some of these views with regard to the effect exerted by consonants, in general, on the duration of vowels that precede them. Special attention will be given to consonant clusters, whether identical or non-identical, and their influence on the duration of the preceding vowel.

The influence of consonant environment upon the duration of preceding vowels has been studied by a great number of scholars in different languages. Earlier, it was reported that both Rousselot (op.cit.) and Sievers (op.cit.) did not explain the effect a geminate consonant has on a preceding vowel. Josselyn (op.cit.), however, followed Rousselot's approach in examining the effect of a geminate on vowel duration in Italian. He found a compensatory shortening of those vowels preceding geminate consonants.

In 1953, Belasco carried out an experiment in which a scale was devised to classify arbitrarily the consonants according to five degrees of 'force of articulation'. The method employed was planned to determine objectively the force of articulation of eighteen different consonants. Only the vowel [ɛ] was used, and it was uttered in a phonetic context which approximated that of a vowel in a typical French accented syllable. It was found that "the duration of the vowel, measured in centiseconds, varied with the consonant that followed: p(14), t(15), k(16), f(18), l(21), n(23), m(25), s(24), ʃ(25), d(26), g(26), b(26), ɲ(28), j(31), v(36), ʒ(37), z(38), r(42)." [p.1016].

Belasco's findings suggest that there is a cause and effect relationship between vowel duration and the voicing of a following consonant. He explains this causal relationship by what he calls the 'force of articulation' of consonants. Belasco maintains that vowel duration varies inversely with the degree of physiological energy required to pronounce the following consonant. His findings are summarised as follows "the anticipation of a consonant requiring a 'strong' force of articulation will tend to shorten the preceding vowel since more of the total energy needed to produce the syllable is concentrated in the consonant. The opposite is true of course when the consonant has a weak force of articulation." [loc.cit.].

The influence of a following consonant on the duration of a tonic vowel in English raises serious linguistic implications and a number of researchers have been led to study the problem cross-linguistically using other languages besides English. For instance, Zimmerman and Sapon (1958) examined the relationship between vowel duration and a following consonant in English and Spanish. They decided to investigate whether "there appears to be any physioacoustic constant governing the duration of stressed vowels, or whether this apparent relationship is primarily a matter of linguistic structure." [p.152]. A list of words was prepared for each language. The lists were read by native speakers of the two languages. The English list consisted of 38 monosyllabic oxytonic words containing vowel [i] followed by all possible consonants. The Spanish list, on the other hand, contained

90 bisyllabic paroxytonic words using five vowels in the tonic position, and the vowels were also followed by all possible consonants.

The choice of the examined materials was strongly criticized by Delattre (1962, p.1142) in that it is unfair to compare bisyllabic paroxytones in Spanish pito/pido and monosyllabic oxytones in English niece/knees. For Delattre a fair comparison would have been Spanish pito/pido with English bitten/bidden so as to have bisyllabic paroxytones in both cases.

On the basis of the samples studied, the findings for both languages indicate that vowels preceding voiced consonants are considerably longer than vowels preceding unvoiced ones. There is, however, an extremely large difference between the range of vowel duration in the two languages. The range in English is 136.0 msec; while it is only 36.1 msec in Spanish. Contrary to the results reported by Belasco (op.cit.), the findings of Zimmerman and Sapon (op.cit.) show that there is no consistent pattern in terms of place and manner of articulation of the following consonant regarding its effect on vowel duration and that their attempt to rank the duration of the vowels in accord with Belasco's notion of 'force of articulation' yielded negative results. These findings are also in agreement with those reported by House (1961) who states that the average duration of vowels varies markedly as a function of the phonetic environment and the primary influence is contributed by the voicing characteristic of the consonant; whereas the manner of production of the consonant shows a smaller effect. The place of consonant production is found to have negligible influence on the duration of a preceding vowel. House remarks that the shortening of vowels preceding voiceless consonants "must similarly be attributed to an articulatory activity arbitrarily imposed by the phonological system of English that calls for the prolongation of vowels before voiced consonant sounds." [p.1177].

House's investigation (op.cit.) was highly favoured by Delattre (1962) who presented new facts and new arguments to reinforce the theory according to which many variations in vowel duration are physiologically conditioned and that those conditioned variations, yet not phonemically learned, are due to factors that are cross-linguistically valid. Apart from stress and tempo, Delattre (ibid.) suggests eight factors that may affect vowel duration in American English. Three of these factors are internal, i.e. in the vowel itself, and the other five are external factors all to be found in the single consonant that follows the vowel. The final conclusion he draws is that variations in vowel length are phonemically learned under only one of the eight factors. He calls this factor the 'abridging/expanding' factor which is commonly known as the lax/tense factor. Variations in vowel length are physiologically conditioned under the seven other factors. He notes that the possibility of conditioning factors being universal, i.e. operating cross-linguistically, is far from negligible.

In his study of consonant gemination in four languages, Delattre (1971) examined vowels preceding geminate and non-geminate consonants. He was curious to know whether the geminates would be preceded by shorter vowels than their single cognates. To his surprise, he found that they were not intrinsically shorter, and assumed that in distinguishing a geminate consonant from a single one, the duration of the preceding vowel is a negligible factor. From his six English subjects, he measured the duration of [i] before [nn] and [n], and of [e] before [ss] and [s]. He obtained the following results:[1]

I've s<u>ee</u>n Nelly.
10(12.5)

I've s<u>ee</u>n Elly.
11(6)

We s<u>ee</u> Nelly.
9.9(7.5)

The r<u>a</u>ce sends.
19(22)

The r<u>a</u>ce ends.
18.2(12)

The r<u>ay</u> sends.
20.1(17)

In the other three languages, namely German, French and Spanish, the results of vowel duration were comparable to those found in English. Vowels were shown to be only slightly shorter before geminates than before single consonants. The results, on the other hand, indicate that the [ɛ] vowels were predominantly longer before a geminate than before a single consonant in the three languages. Vowel

---

[1] Averages (in csec) are followed by the duration of the consonant in parenthesis. (Delattre, ibid., p.43).

[a] had the same tendency in Germany and French. Back vowels were mostly shorter when followed by a geminate in French and Spanish.

Delattre (op.cit.) points out that vowels are not significantly shorter before a geminate consonant than before a non-geminate counterpart, and that the duration of the preceding vowel is not a factor in the perception of consonant gemination. For him these results were unexpected because "vowels are shorter before a voiceless consonant than before a voiced one - an analogical condition with respect to the anticipation of a great effort." [p.112].

Delattre's findings seem to be in total agreement with those reported by McKay (op.cit.). McKay found out that in his materials there were contrasting word pairs in which the vowels before the contrasting plosives were almost exactly equal in length despite the general occurrence of shorter vowels before medial geminate plosives than before the non-geminate ones. For him vowel length in Rembarrnga is considered as non-significant since there is no structural linguistic evidence for phonemic vowel length distinctions.

The duration of the vowel occurring in certain syllabic conditions has also been subjected to some investigation. Peterson and Lehiste (1960), for example, presented a study in which some of the characteristics of duration in English were examined. They contend that if the factors that condition the duration of syllable nuclei can be isolated and described, and if their influence on the syllable nucleus can be determined, the remaining durational differences may then be defined "as intrinsic durational differences associated with each vowel pattern." [p.693].

Peterson and Lehiste (op.cit.) used connected speech for their experiments and the investigated material consisted of two sets of data. The first set contained 1263 words which were selected on the basis of frequency of occurrence. They all had the same structure, consisting of one of fifteen common English syllable nuclei, preceded and followed by a consonant sound. The second set contained 70 minimally different words, including 60 CNC words forming 30 minimal pairs, and 10 disyllabic words constituting five additional minimal pairs. The first set of words was recorded by one speaker in an identical sentence frame, with fixed stress and pitch pattern; whereas the second set was recorded by five speakers who used the same frame sentence and stress and pitch pattern as the speaker for the first set.

The analysed data indicate that the durations of all syllable nuclei in English are significantly affected by the nature of the consonants that follow the syllable nuclei. The influence of initial consonants upon the durations of the syllable nuclei is found to be negligible. The syllable nucleus appears to be generally shorter when followed by a voiceless consonant and longer when followed by a voiced consonant. The data also suggest that plosives are preceded by the shortest syllable nuclei, and nasals have approximately the same influence as voiced plosives. Syllable nuclei are found to be longest before voiced fricatives. (See table 2.6).

Table (2.6) Duration (in csec) of syllable nuclei as a function of the following consonant.

following consonant	duration of short syllable nucleus	duration of long syllable nucleus
p	13.8	18.8
b	20.3	30.7
t	14.7	21.0
d	20.6	31.8
k	14.5	20.0
g	24.3	31.4
m	22.0	31.3
n	21.6	32.2
ŋ	21.8	35.0
f	19.2	26.1
v	23.1	37.4
θ	20.8	26.5
ð	26.0	38.1
s	19.9	26.9
z	26.2	39.0
ʃ	21.2	27.8
ʒ	-	41.0
r	22.6	29.6
l	21.8	29.3
tʃ	14.5	29.8
dʒ	19.1	30.0

(After Peterson and Lehiste, 1960, p.702)

Fintoft (1961) conducted an investigation whose aim was to study the relation between long and short vowel and consonant sounds in Norwegian. The investigation also aimed at examining the effect vowels and consonants have on each other. It is a well-known fact that the length of a syllable is determined both by the vowel and by the following consonant. It also depends on the type of stress and its location. In Norwegian, as well as in many other languages, a short consonant follows a long vowel and a short vowel is followed by a long consonant or a group of consonants. Stressed syllables are always regarded as long syllables, and is comparatively rare that short monosyllables occur without a consonant at the beginning or at the end. It is also more usual to find in Norwegian a single consonant than a combination of consonants both at the beginning and the end of these words.

The test items comprised a number of constructed words which were used in isolation. Each word began and ended with the same vowel or consonant sound. This made it possible to have combinations of consonants and vowels which formed a mixed list of words as well as nonsense syllables. The test items were chosen to have the following syllabic paradigm:

VCV, VCCV, VC, VCC, CVC and CVCC

where the (V) element is [a], [i] or [u], and the (C) element is [f], [s], [v], [m], [n], [l] or [r]. Accordingly, words such as /asa/, /assa/, /sas/ and /sass/ were studied. Although meaningless utterances were included within the test items, the subjects were forced to make a difference between long and short consonants.

Fintoft's final results show that there is a substantial difference between the duration of word-initial vowels followed by short consonants and those followed by long consonants. For instance, the mean duration of vowels followed by short consonants is about 200 msec; while it is only about 100 msec when long consonants exist. This is explained in terms of that the duration of vowels "depends to a great extent on the following consonant." [p.24]. Similarly, vowels in word-medial position are substantially longer before short consonants than before long consonants.

In so far as consonants are concerned, the results denote that the duration of word-medial and word-final voiceless fricatives is significantly longer than that of all other consonants, and that long consonants are characterized by a somewhat longer duration than that of short consonants. In word-initial position, only the duration of [s], [l] and [r] was examined. The results indicate that the average duration of [s] is essentially longer than that of [l] and [r]. At the same time, the average durations of the syllables with short consonants are found to be longer than their corresponding partners with long consonants. The reason behind this is said to be mainly due to the fact that the vowels before short consonants is 1.5 - 2 times longer than the vowels before long consonants; whereas long consonants are only 1.1 - 1.2 times longer than short consonants. This makes it quite obvious that the duration of the vowels in Norwegian is of greater importance to the duration of the syllables than is the duration of the consonants. Furthermore, Fintoft (op.cit.) remarks that "the duration of the vowel is markedly dependent on the following consonant, but seems to be almost independent of the preceding consonant. It seems as if the duration of the vowel is only dependent on the immediately adjacent sounds. The duration of the vowels before 'short' consonants is substantially longer than before 'long' consonants." [p.35].

Another investigation was undertaken by Sharf (1962) to find out whether the duration of intervocalic plosives or the vowels that precede them could act as sufficient clues to listeners in distinguishing between voiceless versus voiced plosives occurring in monosyllabic and bisyllabic words. Most of the words used in the study were members of minimal or near-minimal pairs differing only in the homorganic plosives. Word-pairs like cat/cad, catty/caddy, tack/tag, tacking/tagging, nap/nab and napping/nabbing were studied.

Spectrographic inspection of the studied material shows that the duration ratio of vowels before voiceless and voiced consonants in monosyllabic words is 2:3, and in bisyllabic words it is about 3:4 for vowels preceding [p] and [b], and about 4:5 for vowels preceding [k] and [g]. The average duration of vowels before [d] is 0.9 csec longer than it is before [t]. The inspection also reveals that the closure duration ratio for [b] and [p] is about 2:3, and it is about 3:4 for [g] and [k]. Sharf (ibid.) considers the average difference in duration of 0.9 csec between vowels preceding [t] and [d] as a sufficient cue in distinguishing them.

Schwartz (1969, 1970), on the other hand, was interested in studying the influence of vowel environment upon the duration of fricative consonants. He devised experiments to consider the possibility of a different effect of vowel environment on the duration of [s] and [ʃ]. Bisyllabic nonsense words of the pattern (VCV) were constructed in which the (V) element stands for either [i] or [a] in both positions.

The analysis of the data shows that [s] and [ʃ] in the environment of [i] appears to be longer than those in the environment of [a]. According to this finding, it is hypothesized that [i] has a lengthening effect upon both [s] and [ʃ]; while [a] has the reverse effect. This suggests that the durational differences of [s] and [ʃ] are the result of an influence of the final vowel and not of the initial one. Schwartz (1969) reports that "this interpretation is drawn from the finding that the fricatives in the environment of /a - i/ and /i - i/ are significantly longer than those in the environment of /a - a/ and /i - a/ - whereas no significant fricative-duration differences exist between the /i - a/ and /a -a/ environments and between the /i - i/ and /a - i/ environments." [p.481]. It is speculated that the shorter duration before [a] represents an earlier release of the consonantal constriction which may be viewed as an anticipatory response to the greater spatial distances which must be covered by the articulators when proceeding from the fricative to [a] than when they proceed from the fricative to [i].

In addition to the stimulus items examined in the first investigation, Schwartz (1970) inserted three plosives after [s] to construct bisyllabic words having the pattern (VCCV) where the first (C) element was always [s] and the second was either [p], [t] or [k]. This procedure was devised to examine the duration of [s] in /s/-plosive combinations.

The analyzed results indicate that there is a significantly shorter duration for [s] before [p] than before [t] and [k]. These findings suggest the existence of a coarticulation between the plosives and the preceding fricative. Schwartz (op.cit.) states that "the significantly shorter duration of [s] before [p] when compared to [t] and [k] may thus be interpreted as a manifestation of a tendency to keep /s/-plosive durations constant in ongoing speech." [p.1144].

Contrary to those results reported by Schwartz (1969), Umeda (1977) conducted an experiment in which it was shown that the sequential effect of vowels on consonants was not seen in data from an extensive reading. The reason is probably because the effect was overshadowed by other higher-level factors when the speech material increased in complexity. Umeda (ibid.) states that consonant durations are a function of six different factors. The factors are:

(1) position of the consonant in the word;
(2) its relation to lexical stress and morpheme boundary (if any) within the word;
(3) whether it is in the postpausal position;
(4) whether it is in the prepausal position;
(5) content-function difference of the word; and
(6) effect of adjacent consonant both inside the word and across the word boundary.

Two other major factors that determine the duration of a consonant in a vowel environment are also mentioned. The two factors are the position of the word-boundary and the position of stress. The first factor is related to whether the consonant is immediately after a word boundary, before the boundary, or inside a word; whereas the second factor is related to whether or not the consonant is at the head of a stressed syllable. The studied data indicate that consonants show quite regular durational behaviour when they are surrounded by vowels, but when adjacent to another consonant their durational patterns become complex. Generally, a consonant is longer when a voiced consonant precedes it than when a voiceless consonant does so. [s] is found to be an exception among voiceless consonants; when it precedes [ð], the latter becomes long.

Following an earlier experiment carried out by House and Fairbanks (1953), Chen (1970) studied extensively the effect of following consonants on the preceding vowel length in English, French, Korean and Russian. He examined the variations of vowel length as a function of the [± voice] feature of the following consonant. According to him such vowel durational differences are either language-specific, i.e. "primarily a matter of linguistic structure", or language-universal, i.e. "conditioned by an inherent physiological feature of articulation." [p.130]. The measurements obtained from 376 spectrograms taken from recordings of the four languages under investigation led Chen to conclude that variations of vowel duration as a function of the voicing of the following consonant is presumably a language-universal phenomenon; while the extent to which an adjacent voiced or voiceless consonant affects its preceding vowel duration is language-specific. It is determined by the phonological structure of the language. The word-list, of each language, was read by native speakers and care was taken to assure identical accentual patterns for each member of the minimal pair so as to "eliminate variability in duration due to suprasegmental features." [pp.130-131].

Chen's final results illustrate that in all four languages a vowel is invariably longer before a voiced consonant than before a voiceless one. (See table 2.7). For him such invariable variations of vowel duration depending on the voicing of the following consonant "can hardly be regarded as accidental." [p.135].

Based on his own findings and those reported by other researchers (e.g. Zimmerman and Sapon, op.cit., Fintoft, op.cit.), Chen strongly concludes that the variability of vowel duration as a function of the [± voice] feature of the following consonant is language-universal. These findings confirm that vowels generally tend to become longer before voiced consonants and shorter before voiceless ones. However, he remarks that the voicing of the adjacent consonant influences its preceding vowel to different degrees in different languages. He states that vowel duration in English varies "rather drastically depending on the following consonant." [p.138]. The vowel preceding a voiceless consonant is only 0.61 or less than 2/3 of that preceding a voiced consonant. The case seems different in the other three languages. For instance, in French the ratio is 0.87 which means that the shorter vowel is only 13% less in duration than its longer cognate.

The examined data show that the occlusion time before the release of voiceless plosives is appreciably longer than that before the release of voiced ones. The vowel duration is shown to vary inversely with the closure time of the following consonant. This is interpreted that the longer the closure time for voiceless consonant, the shorter the vowel length, and the other way round. As far as this point is concerned, Chen reviews Kozhevnikov and Chistovich's (1965) consideration of the variability of vowel duration in question as "the assumed effect of the compensation of differences in the duration of closure by differences in the duration of the adjacent vowels; however, this compensation was not full." [p.107].

Table (2.7) A cross-linguistic view of vowel length variation as a function of the voicing of the following consonants.

	vowel length (in msec)			
	before voiceless consonants	before voiced consonants	mean difference	ratio
English	146	238	92	0.61
French	354	407	53	0.87
Korean	91	119	28	0.78
Russian	131	160	29	0.82

(After Chen, 1970, p.138)

Nooteboom and Slis (1972) investigated the phonetic feature of vowel length in Dutch. They noted that their results confirm that consonants following a short vowel have a consistently longer duration than consonants following a long vowel. They found evidence which suggests that in nearly all cases the [p] consonants following the phonetically short vowels have a longer duration than those following the phonetically long vowels. The duration of a [p] following a short vowel, for instance, was about 10 msec longer than that of a [p] following a long vowel. They believe that "there is evidence from the literature that this longer duration is the result of a stronger closing command after short vowels." [p.315]. They contend that their findings are in harmony with those reported by other researchers, such as Fischer-Jørgensen (1969) who studied extensively the articulatory, acoustic and perceptual differences between consonants following German long and short vowels, especially in view of the perceivable difference between close and loose contact.

Earlier, Slis and Cohen (1969) had reported that vowels preceding voiced consonants are generally longer than those preceding voiceless ones. The acoustic measurements show that if the consonants differ in voice character only, a particular vowel preceding a voiced consonant is always longer than when followed by a voiceless consonant within the same context. On the other hand, the perceptual tests indicate opposite results. Vowels preceding voiced consonants turn out to be 25 msec longer than those preceding voiceless ones. The objective of these tests was to find out "whether a perpectual influence on the length of the vowel could be established depending on the voice character of the following consonant." [p.89].

It is a well-known fact that stressed vowels are shorter before voiceless consonants than before voiced ones, and that stressed vowels are shorter before an unstressed syllable in a bisyllabic word than in a monosyllabic word. (House and Fairbanks, 1953; Denes, 1955; House, 1961; Peterson and Lehiste, 1960; Lehiste, 1972). Klatt (1973) designed a set of test materials to study the interaction between these two rules and examine the validity of the 'percentage change model' by inspecting one situation where two duration rules interact. He constructed a list of 80 words half of which were monosyllabic words and the other half were bisyllabic words. The test words were spoken in frame sentence "in order to produce speaking rates more nearly in line with conversational speech and to avoid effects due to prepausal lengthening in utterance-final position." [p.1102].

The final results suggest that vowels become strongly incompressible beyond a certain amount of shortening and that the stressed vowels tend to be shorter in the two-syllable words than in the one-syllable words. There is about twice as much shortening of a vowel when a second syllable is added to a word with voiced postvocalic consonant.

Elsewhere, Klatt (1976) remarks that rules involving segmental duration may add and subtract increments to durations, multiply durations by scale factors, or operate by more complex formulae. He states that "vowels shortened by one rule behaved as if they were now somewhat incompressible in the sense that both an additive change model and the percentage-change model predict a duration that is too small." [p.1215].

Recently, Port et al. (1980) have conducted two experiments to explore the effect of changing inherent consonant durations on adjacent vowel in two languages. More specifically, they studied the effects of changes in voicing and manner of 'apical obstruents' both in Arabic and Japanese.

In their Arabic experiment they examined timing in words with [t], [d] and [r] following the stressed vowel when spoken at three distinct speaking rates. The stressed vowel was either the short [a] or the long [aa]. The investigators wished to know "if the voicing or manner change would affect the preceding vowel duration and whether there would be any inverse relation between the vowel and the post-stress consonant constriction at the three speaking tempos." [p.237]. The results of this experiment give minimal evidence of temporal compensation in the studied Arabic syllables. There is no evidence that vowel lengthening for voicing of a subsequent consonant is related to the shortening of the consonant itself. In this respect, Arabic differs from English. The results indicate that the difference in duration between [d] and [t] is very small, which implies that (VC) duration is greater for the syllables with [d] than with [t].

The only evidence of temporal compensation is the lengthening of the vowel before the very short [r].

The Japanese experiment was, on the other hand, devised to explore the question "as to how successful Japanese speakers will be at keeping the CV unit constant in duration if either two very long segments or two very short segments combine in a single syllable." [p.241]. For this experiment Port et al. (op.cit.) employed bisyllabic words containing the longest vowel [a], the shortest vowel [u], and four apical consonants ranging from the very long [s] to the very short flapped [r].

In contrast to those results reported by Han (op.cit.) the findings of this experiment do not support Han's specific proposal that "the domain of temporal compensation is strictly within the CV unit." [p.244]. Temporal compensation appears to extend across several syllables so that two-syllable units have a nearly constant duration despite the fact that great changes occur in the duration of individual consonants and vowels. Thus, vowels vary inversely with adjacent consonants, and consonants vary inversely with adjacent vowels and even with consonants in following syllables.

Comparison between the results of the two experiments show that whereas in Japanese the preceding and following vowels compensate for the changes in consonant duration to about the same degree, in Arabic, on the other hand, the following vowel is unaffected by the consonant, and only the preceding vowel varies as a function of the consonant. Port et al. (op.cit.) state that Arabic and Japanese appear to behave differently from each other in similar contexts implying that these effects are not physiological necessities. They represent timing behaviour that must be learned by speakers.

The findings of Port et al. were later confirmed by Fledge and Port (1981) who also showed that in Arabic the duration of long vowels preceding voiced plosives, namely [d] and [g], was not significantly longer than vowels preceding voiceless plosives, namely [t] and [k]. This result seems to violate the claimed universality of the plosive voicing effect on preceding vowel duration suggested by Chen (op.cit.). Fledge and Port concluded on the basis of their findings that "this phonetic context effect on vowel duration may not be a phonetic universal as is often supposed." [p.131].

PART TWO

PERCEPTION AND PRODUCTION OF GEMINATION

CHAPTER THREE

INITIAL MEASUREMENTS AND ACOUSTIC CHARACTERISTICS OF CONSONANTS IN ISOLATE AND CONTEXTUAL UTTERANCES

3.1 Introduction and Aims

Relevant studies on gemination for a number of different languages, as explained in Part One of this study, have shown that geminate consonants are longer than their single partners. For instance, Delattre's (1971) spectrographic and cineradiographic investigation of the phenomenon of gemination in four languages showed that the duration of a consonant is considered as a major factor in the linguistic functioning of gemination and that the duration of the preceding vowel is shown to be a negligible factor in the perception of consonant gemination, and vowels are not essentially shorter before a geminate consonant than before a non-geminate one. In their study of consonant duration in two different languages, Lehiste et al. (1973) confirmed that the duration of a single consonant is significantly shorter than that of its geminate counterpart, and that differences between single and geminate consonants in word-medial positions are strongly established. On the other hand, McKay's (1980) spectrographic investigation of word-medial plosive geminates in Rembarrnga indicated that the occlusion durations for geminate and non-geminate plosives showed that the geminates are considerably longer than the non-geminates. In contrast to those results reported by Delattre (op.cit.), McKay found that in some cases vowels are insignificantly shorter before the geminate consonants than before the non-geminate ones.

Earlier, in section 1.3, it was mentioned that Arabic is a language which exploits gemination regularly and extensively in morphological processes, and that gemination is frequently encountered in the systems of derivation and inflection in its structure. Despite the large literature on Arabic dialects, one may contend that the whole matter of gemination in Arabic needs more detailed investigation. There is by no manner of means any complete or satisfactory instrumental study or structural description of the phenomenon of gemination in any variety of the Arabic language. Obrecht's (1965) perceptual investigation of Arabic geminate consonants, for example, was limited in scope. He only examined contrasts in three phonetic categories and suggested that duration is regarded as the determining cue factor to distinguish a single consonant from its geminate cognate in Arabic. Hassan's (1981) study, on the other hand, was mainly concerned with investigating experimentally vowel duration in Iraqi Arabic. His test stimuli to examine the durational differences between vowels preceding geminate consonants and those preceding non-geminate consonants were limited in number. However, his results indicated that geminate consonants are significantly longer than the non-geminate ones, and that vowels are essentially longer before single consonants than before their geminate partners.

The experiment to be described below is one of a series aimed at investigating the significance of gemination in the Arabic language in general and in I.C. Arabic in particular. It also aims at investigating durational differences between single and geminate consonants occurring initially and medially in words pronounced in isolation and in words pronounced in a carrier sentence so as to determine whether or not duration can be considered as a major factor in the contrast between single and geminate consonants in I.C. Arabic.

3.2 Experimental Procedure

3.2.1 Selection of Stimuli and Recording Technique

In this experiment two oppositions involving different contoid types were chosen; [s] versus [ss] and [d] versus [dd], each in word-initial and word-medial position. The words /sabit/ 'Saturday' versus /ssabit/ 'the Saturday', /ħasan/ 'Ḥasan' - a name of a male person - versus /ħassan/ 'he improved', /darub/ 'road, street' versus /ddarub/ 'the road, the street', and /badal/ 'substitute (n)' versus /baddal/ 'he altered' were selected for these oppositions. Therefore, discrimination between single and geminate fricatives and plosives was involved. All the test words selected for this experiment bear primary stress on their first syllables. A carrier sentence, viz. /ˈkitbi——ˈsit marˈraat/ 'Write —— six times', was constructed in which these words were employed. Each isolate word or carrier sentence was repeated twice.

A list, making a total of 160 items, was recorded in the recording studio of the Phonetics Laboratory on high quality Ampex tape (type 1200). The tape recording speed was 7½"/sec, and the microphone used was AKG (type D202E1). The distance between the microphone and the mouth of the speaker was 30 cm. The recorded items were distributed on 10 individual runs. Each run comprised 16 words, i.e. 2 tokens for each word. The first 5 runs contained the words spoken in isolation; the other 5 contained the words uttered in a carrier sentence. In each run the recorded items were randomized anew with the help of a microcomputer (type Commodore PET 2001 series). The computer was placed in front of the speaker inside the recording studio while the recordings were made. It took 2 seconds between the appearance of one word and the next on the screen of the computer while saying the words in isolation, and 3 seconds in the case of pronouncing them in a carrier sentence. The reason behind arranging the test words in random order was to avoid monotony and to make the enunciation of the sentences seem natural. The dummy word /kitaab/ 'book' was inserted at the beginning of each run.

3.2.2 Subjects

All the studio recordings were made by the experimenter as his own subject. He was born in Basrah, one of the biggest cities in the southern region of Iraq. He had lived there almost all his life, and had been educated there, except for the years between 1960 and 1963 where he lived in Baghdad to complete his higher education. He speaks a typical I.C. Arabic with a dominating accent of the city of Basrah. He has essentially normal speech and hearing.

3.2.3 Instrumental Set-up

Intensity level traces were obtained by playing back the recorded materials on a stereo tape-recorder (type Revox A77). The output of the tape-recorder was on channel 1. The gain was set at 7 and the left/right balance was set at 12 o'clock. The output signals from the tape-recorder were fed into an intensity meter (made by Frøkjaer-Jensen Electronics). Three types of intensity level traces were obtained; high pass (H.P. for short) filtered at 3.9 kHz (log scale), H.P. filtered at 500 Hz (log scale) and low pass (L.P. for short) filtered at 500 Hz (linear scale). The output signals from the intensity meter were graphically recorded on a mingograph (type Mingograf 803 - Siemens Elema) and, thus, mingograms were made of the audio recordings. The speed of the mingograph paper was 5"/sec. Three different traces for three separate channels were marked on the paper. Channel 1 displays the trace of the H.P. filter (3.9 kHz) output, channel 2 displays the output of H.P. 500 Hz and channel 3 that of the 500 Hz L.P. filter. From the H.P. traces it was possible to identify fricative consonants. Vowels were easily identifiable from the L.P. traces as well as from the H.P. traces filtered at 500 Hz. Plosives, on the other hand, were identifiable at the baseline of almost all types of traces. See Fig. 3.1 for the block diagram used

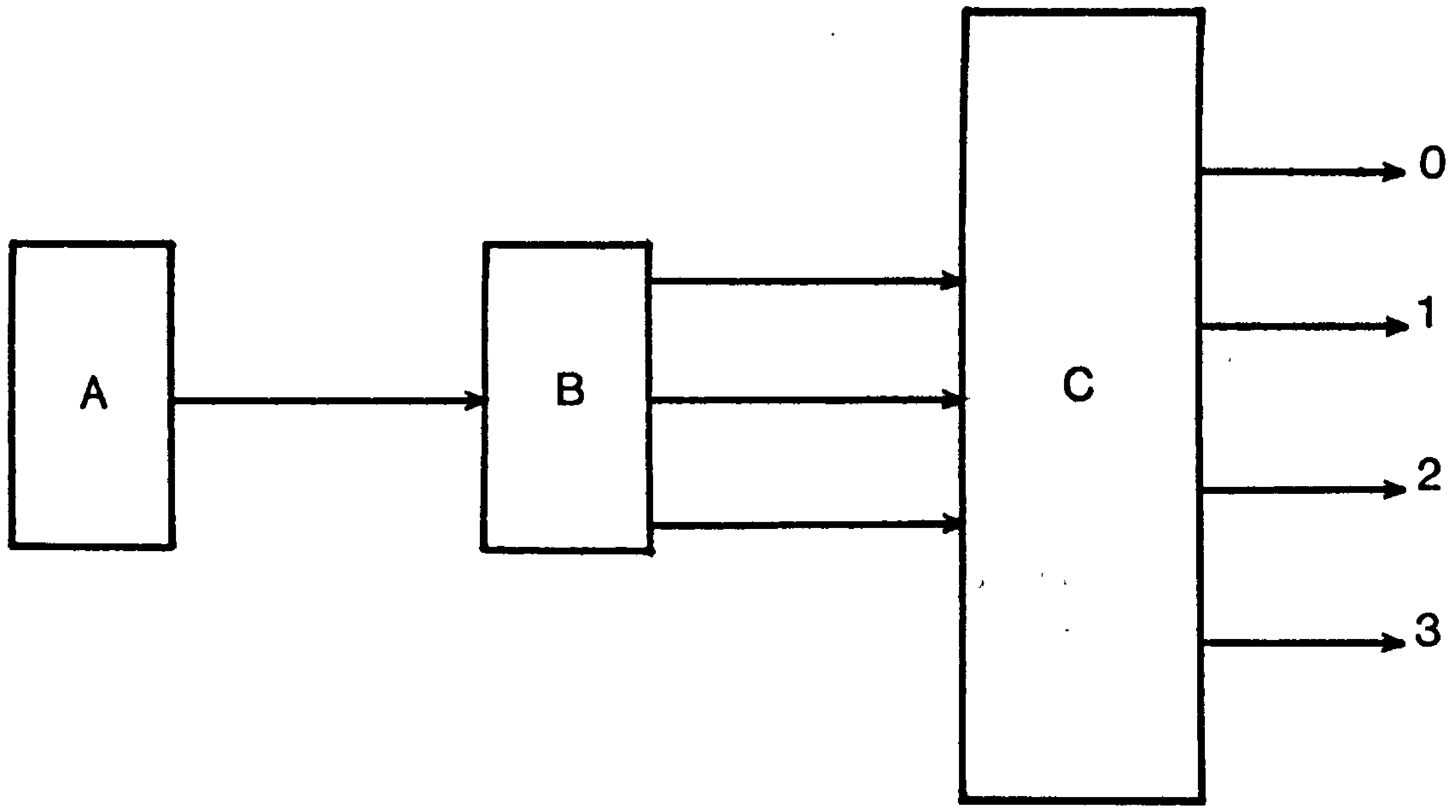


Fig.(3.1) Block diagram of instrumental set-up used in the acoustic experiment of this study

Fig. (3.1)

Description:

A	Revox tape-recorder (type A77).
B	Intensity meter (F-J Electronics).
C	Mingograph (type Mingograf - 803/Siemens Elema).

Mingograph traces:

O	Timer
1	Channel 1 showing intensity level H.P. filtered at 3.9 kHz (gain set at 7¼).
2	Channel 2 showing intensity level H.P. filtered at 500 Hz (gain set at 7).
3	Channel 3 showing intensity level L.P. filtered at 500 Hz (gain set at 7½).

in this experiment.

In addition, the 700 spectrograph was exploited as a supporting technique to help solve some inevitable problems which would be encountered while making our segmentation. Therefore, the speeds of the two machines, namely the mingograph and the spectrograph, were set to match each other, i.e. about 129 mm/sec.

When wide-band spectrograms were made for some of the studied items, the output of the Revox tape-recorder was set at 8+ channel 1 and the left/right balance was set at 12 o'clock. The input to the spectrograph was recorded at -OdB on the V.U. meter for the reference tone; whereas the input level of the spectrograph was set at 5. When producing the required spectrograms, the scan level of the spectrograph was always set at 1½ so that the V.U. meter was zero for the strongest items which would be considered as the norm for the other recorded items. This procedure was so done in order to get consistency of segmentation made on mingograms showing intensity level traces and another made on spectrograms for the same stimuli.

3.2.4 Segmentation and Measurements Criteria

Segmentation of speech sounds very often constitute a major difficulty confronting the researchers. Peterson and Lehiste (1960), for instance, reported that segmentation "has long been and continues to be a major problem in speech analysis." [p.694]. This is largely due to the structure of the vocal system which is considered as a complex network with many coupled or interacting components. The physiological complexity is also reflected in the acoustical structure of the produced sounds, for "the acoustical signal does not show a simple one-to-one correspondence between the physiological and the acoustical patterns." (Peterson, 1955, p.418). Peterson illustrated this fact by stating that an intervocalic voiceless plosive may involve a considerable period of silence during which major portions of the vocal mechanism are in motion.

Phoneticians have for a long time realized the difficulty of segmenting the continuous speech signal into a succession of discrete segments. Abercrombie (1965), for instance, was well aware of this fact when he remarked that the concept of the speech sound as a stable posture of the vocal organs is a fiction. He believes that a parametric approach is more useful, in as far as the linguistic sciences are concerned, than a segmental approach, and that the segmental approach is now giving way to the parametric approach. He states that "The division of speech into phoneme-representing segments represents a division at right-angles to the time axis, whereas the division into parameters is a division parallel to the time axis." [p.123]. Laver (1970), on the other hand, seems to be in total agreement with Abercrombie's contentions. He claims that the parametric approach looks upon the physiological system for speech as "a single, complex system in which the continuous interacting activities of the various linked components are intricately coordinated in time." [p.53]. According to him, the parametric approach leads fairly directly to an interest in the neural control systems capable of coordinating the complex movements of the vocal organs during speech production.[1] Raphael et al. (1980) state that generally there is no acoustic criterion that directly divides the speech stream into segments corresponding in size to consonants and vowels. This, they believe, is due to the fact that the processes of articulation, and especially of coarticulation, cause the phonetic information to be overlapped as it is encoded in the acoustic signal and, consequently, "the segments of the signal, however defined, cannot be matched straightforwardly to the segments of the phonetic message." [p.297]. Earlier, Liberman et al. (1967) pointed to the same fact. They contend that if one examines the acoustic cues more generally, one will find that "successive phonemes are most commonly merged in the sound stream." [p.441]. This contention led them to suggest that segmentation exhibits a complex relation between linguistic structure or perception, on the one hand, and the sound stream on the other. Nevertheless, the boundaries of certain speech sounds may unambiguously be defined. House and Fairbanks (1953) claimed that "the identification of the beginning and end of a vowel surrounded by consonants is an arbitrary act that is both difficult and artificial. Location of these points was aided by the relative clarity with which they are shown in sound spectrograms." [p.107].

Investigators of speech sounds have employed various advanced techniques for the purpose of segmenting speech

---

1. However, Roach and Roach (1983) have recently shown that it is possible to segment continuous speech into discrete acoustic segments automatically with sufficiently sophisticated rules, with a reasonable degree of accuracy.

utterances. They believe that for purposes of automatic speech recognition, it is obviously necessary to use some procedure for segmentation or quantization. (Preston and Lehiste, op.cit.). They also believe that there are circumstances in which it is feasible to signal out, without ambiguity, the points where a certain speech sound initiates and/or terminates within a higher-level phonological unit. They claim that some segmental boundaries can be established with considerable precision from acoustic records. Meanwhile, there are other circumstances where it is very difficult to specify the points of segmentation. Still, one can say that the problem is not so much in measuring as in determining the points at which to perform the measurements. (Lehiste, 1970).

In our present experiment we have relied on the variations indicated by the intensity level traces for the purpose of determining segmental boundaries and for measuring their acoustic durations on the obtained mingograms. An attempt was made to specify the relevant procedures for segmentation to make durational measurements easier. These procedures can by no means be considered universal. The L.P. intensity level traces filtered at 500 Hz were used for making our acoustic segmentation and measurements. The H.P. intensity level traces filtered at 500 Hz were found to be a useful support to which we had recourse when making our segmentation of the consonants, and especially of the fricatives. The durational measurements were made between points in time at which the intensity level traces crossed a threshold value of intensity level which was arbitrarily set at 26 dB above the baseline, i.e. OdB. The 26dB level was accepted as a reference level and was employed for all the measurements of the vowels as well as the consonants. The choice of this arbitrary threshold value was found to be most suitable for making our measurements partly because it allows durational measurements for those consonants and unstressed vowels whose peaks might not go higher than 26dB, and partly because it brings about almost perfect correspondence between durations made on mingograms and those made on some spectrograms. See Fig. 3.2.

In cases where the intensity level traces did not reach the reference level, the segmental boundaries were determined by looking at the points where the intensity level traces did not fall as far as the arbitrary threshold value, i.e. the reference level. The points at which sudden changes of either rises or falls of the intensity level traces were considered as points representing the beginning and the end of a vowel. These rises and falls correspond to similar patterns on spectrograms. For example, a sudden and steady rise in the intensity level trace to a higher level usually corresponds to the commencement of those patterns which are heavily marked for higher formants characterizing vowel sounds on spectrograms. These points were marked, where applicable, by the two letters (c) and (d) on the mingograms. Aiming at getting accuracy in our measurements, vertical lines were drawn from these points down to the reference level. The two letters (a) and (b) were meant to represent the beginning and the end of a vocalic segment on the reference level. See Fig. 3.3.

While carrying out the durational measurements, whether they were in words spoken in isolation or in a carrier sentence, the starting and the ending points of the entire word or the carrier sentence were also determined. The point at which the intensity level traces departed from the baseline was regarded as the initial point of the word or the carrier sentence, and the point at which they came to meet again was regarded as the terminal point.

Despite the difficulty usually encountered in demarcating the beginning of a plosive consonant occurring word-initially, the following criteria were adopted while making our segmentation. The point where there was an abrupt rise of the intensity level trace, indicating the initiation of a following vowel, was taken to represent the end of the plosive sound. The beginning of the release of the plosive was, then, specified from a point where a short single spike started. Peterson and Lehiste (op.cit.) noticed the same observation while they were making their measurements on the spectrograms. They remarked that "after voiced initial plosives, the period of aspiration was absent, but the period of frication following the spike was usually more prominent than in the case of voiceless plosives." [pp.694-695]. The spiky curve was clearly distinguishable on the L.P. intensity level trace, and particularly in the case of the geminate [dd]. See Fig. 3.4. In our measurements the spiky curve of word-initial [d] and [dd] was included. Similarly, the aspiration phase of the voiceless plosive [t] was included in the measurements when it occurred word-finally, as in

/sabit/ and /ssabit/. See Fig. 3.2.

Generally speaking, the durational measurements of the fricative consonants were easy to carry out. The boundaries of the vocalic segments, whether they were preceded or followed by a voiceless fricative such as the sibilants [s] and [ss], or the pharyngeal [ħ], were very clear. See Fig. 3.5. Contrary to the plosive consonants, the fricatives occurring initially in a word were characterized by having their intensity level traces high above the baseline on both H.P. intensity level traces; whereas they took the same level as that of the baseline on the L.P. traces. No spiky curves were seen to initiate voiceless fricatives in word-initial position. See Fig. 3.2.

Other problems in segmentation were encountered while determining the boundaries between a vowel and a word-final nasal or lateral consonant, as in the words /ħasan/, /ħassan/, /badal/ and /baddal/. The intensity level traces of the second vowel [a] together with those of the following nasal or lateral seemed as if they were to form a succession of closely linked peaks that ended in a gradual descending curve. In order to set this problem in clear perspective, the first sharp fall taking place within this succession of peaks was considered as the ending point of the preceding vowel as well as the starting point of the following nasal or lateral consonant. The lateral consonant was seen to continue its sharp successive spiky traces, in their descending scale, on the H.P. traces filtered at 500 Hz; while they died away in a rapid steep fall in the case of the nasal

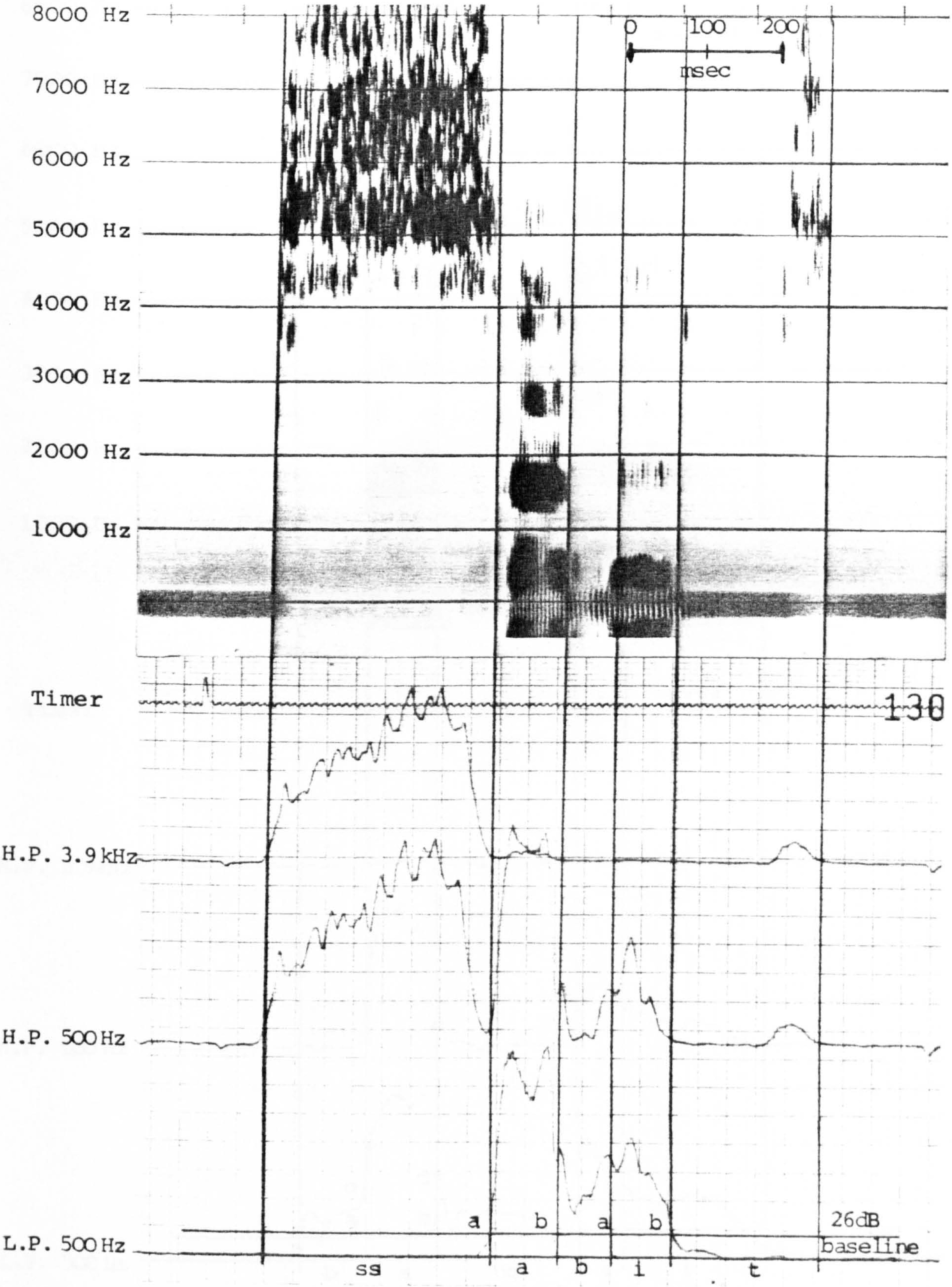


Fig. (3.2) Spectrogram and mingogram of the word /ssabit/ said in isolation. On the mingogram, a and b are points in time where the L.P. intensity level trace crosses the arbitrary threshold value of 26dB. The distance between these two points represent the duration of the two vowels [a] and [i].

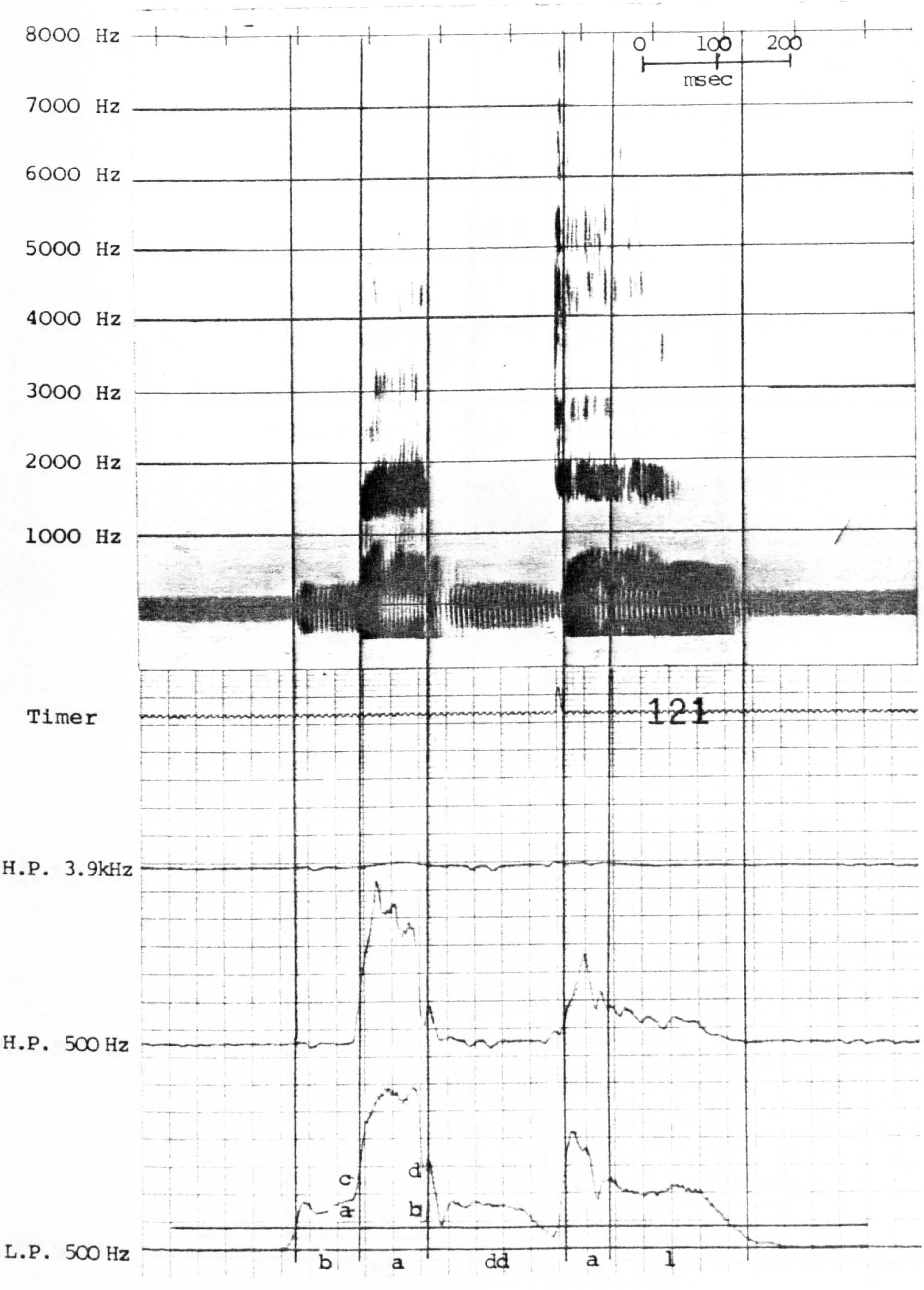


Fig. (3.3) Spectrogram and mingogram of the word/baddal/ said in isolation. On the mingogram, c and d are points in time where sudden changes of slope occur in which the intensity level trace does not reach the threshold value of 26 dB.

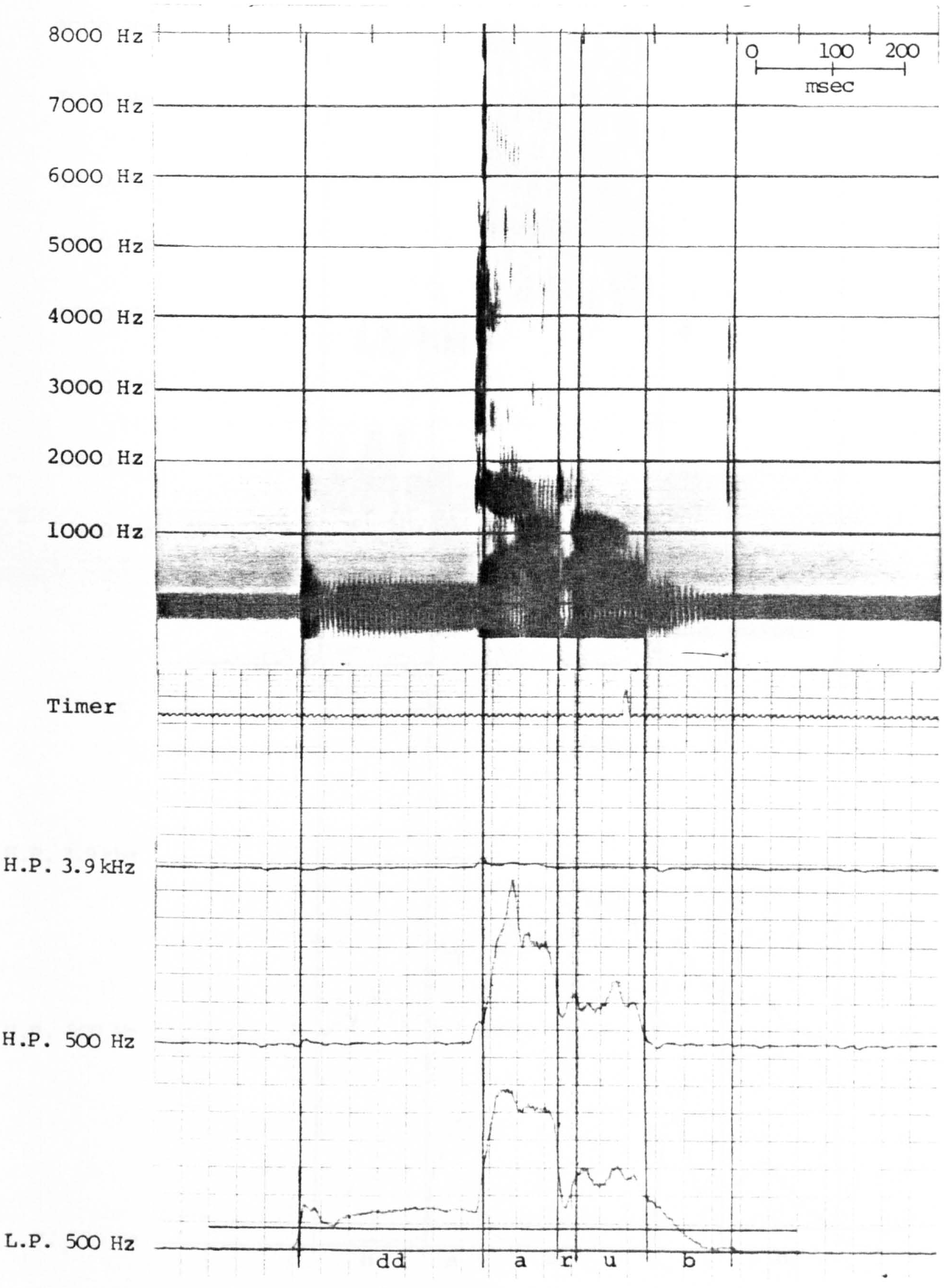


Fig. (3.4) Spectrogram and mingogram of the word /ddarub/ said in isolation

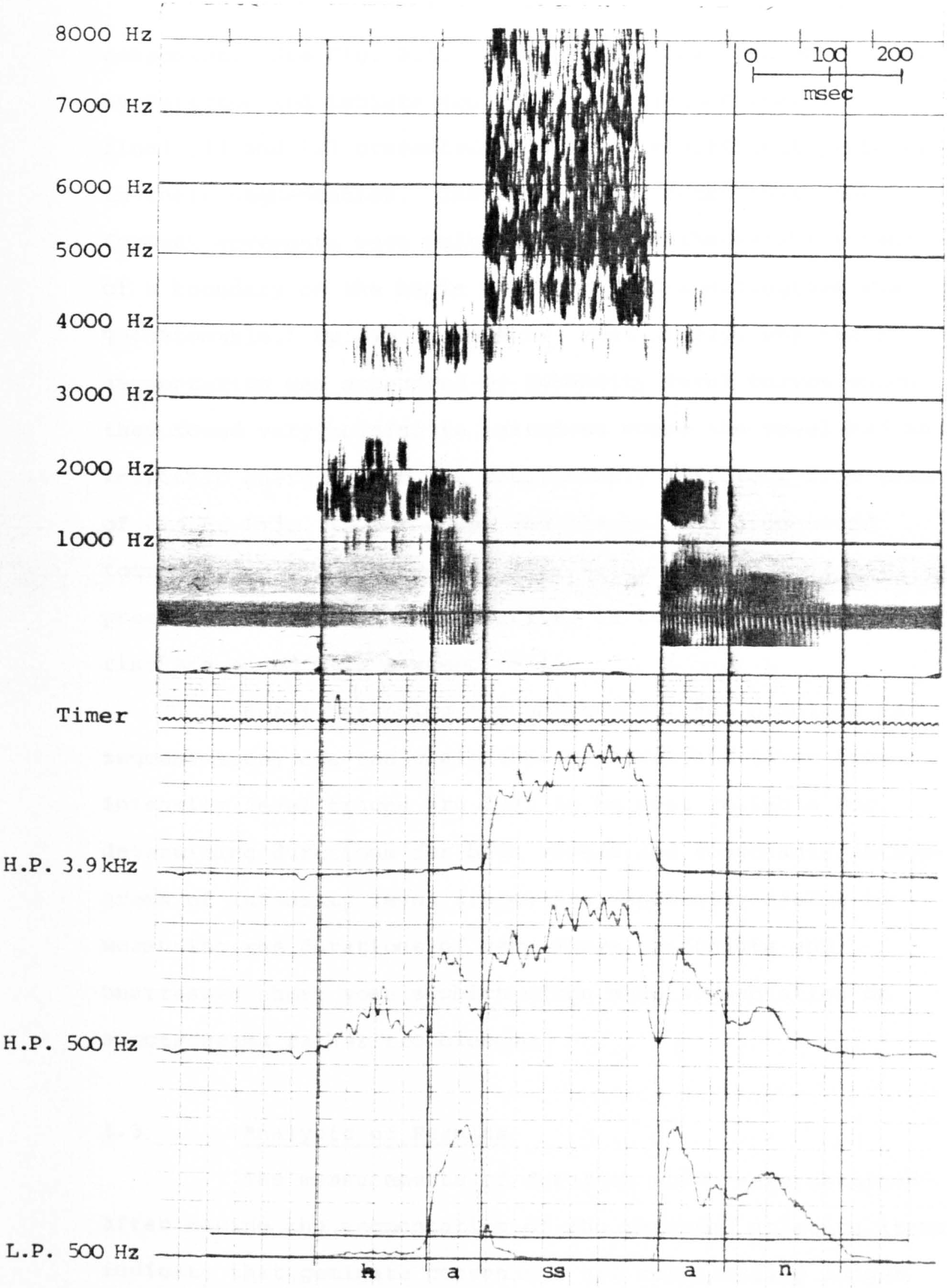


Fig. (3.5) Spectrogram and mingogram of the word /ħassan/ said in isolation consonant. See Fig. 3.3. This problem was also encountered by Peterson and Lehiste (op.cit.) who claimed that word-final [l] and [r] presented particularly difficult problems in their segmentation. They frequently found that "the formant movements were quite smooth, and the establishment of a boundary on the basis of broad-band spectrograms was questionable." [p.698]. This is, most likely, why their segmentation was supported by intensity level curves which they found very helpful in instances where the vowel had an intrinsic energy which was considerably different from that of [l] or [r]. In our data, the flapped [r] pronounced intervocalically produced similar traces to those of a preceding or following vowel, i.e. it began with a sharp rise and ended with a steep fall.

Notwithstanding the above-mentioned problems in segmentation, one can claim that measurements taken from intensity level traces are felt to be most reliable for determining durations for both vowels and consonants. Mingograms of intensity level traces are especially useful in measuring the durations of very short consonants and unstressed short vowels which often make segmentation on spectrograms rather troublesome.

3.3 Analysis of Results

The measurements of duration which were obtained after making the segmentation of the recorded stimulus items indicate that geminate consonants are considerably longer than their non-geminate partners whether they are spoken in isolation or in a carrier sentence.

Based on the measurements for 40 words, where the durations of initial [s] and [ss] were calculated, durational differences between single and geminate consonants are clear. They are shown in tables 3.1 and 3.2 where the mean duration of [s] is 187 msec when pronounced in words spoken in isolation and 138 msec when pronounced in words spoken within a carrier sentence; whereas it is 325 msec for [ss] when pronounced in words uttered in isolation and 219 msec when uttered within a carrier sentence. In the same way, measurements were also made for 40 words in which [s] and [ss] were used word-medially, and the durational differences showed a similar trend to that of word-initial [s] and [ss]. They are shown in tables 3.3 and 3.4 where the mean duration of [s] is 121 msec when pronounced in words spoken in isolation and 102 msec when pronounced in words used in a carrier sentence. [ss], on the other hand, has the mean duration of 254 msec when uttered in words said in isolation and 209 msec when uttered in words said in a carrier sentence.

Since the aim of this experiment was to investigate durational contrasts between geminate and non-geminate consonants in two oppositions involving different contoid types in I.C. Arabic, measurements were also made for the two plosives [d] and [dd], and 80 words in which these two consonants contrasted, either initially or medially, were segmented and the closure durations of the two contrasting sounds were measured. Albeit it is not easy to measure accurately the duration of a plosive consonant occurring initially in words pronounced in isolation, an attempt was made[1] to measure the closure duration of initial [d] and [dd] pronounced both in words spoken in isolation and in words said in a carrier sentence. Considerable durational differences were noticed, and the geminate [dd] was found to be several times longer than its non-geminate counterpart [d], especially when the two plosives occurred medially in words uttered in a carrier sentence. The durational differences are shown in tables 3.5–3.8, where tables 3.5 and 3.6 indicate that the mean duration of word-initial [d] is 73 msec when pronounced in words said in isolation and 46 msec when pronounced in words used in a carrier sentence. [dd] has the mean duration of 224 msec when uttered in words spoken in isolation and 156 msec when uttered in words used in a carrier sentence. In word-medial position, [d] has the mean duration of 29 msec when pronounced in words spoken in isolation and 24 msec when pronounced in words said in a carrier sentence; whereas [dd] has the mean duration of 153 msec when pronounced in words said in isolation and 142 msec when pronounced in words spoken in a carrier sentence.

In addition to the measurements made for the segmental units individually, i.e. [s] versus [ss] and [d] versus [dd], either in word-initial or word-medial positions, measurements for the other sound components forming the whole target word, whether enunciated in isolation or within a carrier sentence, were also made. The durational measurements of these target words are shown in tables 3.9 - 3.24.

---

[1] See section 3.2.4 for details.

Tables 3.9 - 3.12 indicate that the duration of the geminate consonant [ss] in the word /ssabit/ is considerably longer than that of its single cognate [s] in the word /sabit/. The difference in duration between these two cognate sounds helped, to a great extent, in the resultant difference of the ultimate durations of the two words whether uttered in isolation or in a carrier sentence. The total mean durations of the word /sabit/, as shown in these tables, are 655 msec and 402 msec; whereas the word /ssabit/ has the total mean durations of 755 msec and 486 msec, respectively. The vowel sound [a] is almost of the same durational value in both words, especially when the words /sabit/ and /ssabit/ were pronounced in a carrier sentence, see tables 3.11 and 3.12. The word-medial plosive consonant [b] has the durational value of 45 msec when the word /sabit/ was said in isolation or in a carrier sentence; while the difference is clear in the case of the word /ssabit/, see tables 3.10 and 3.12. The durational difference in the enunciation of the voiceless denti-alveolar plosive [t] occurring finally in the two words is mainly due to the result of having a longer aspiration following the [t] of the word /sabit/ than in that of the word /ssabit/. This phonetic phenomenon was observable in almost all measurements made on the mingograms for both words said in isolation. On the other hand, the long aspiration phase that accompanied word-final [t] in the two words disappeared when the same words were spoken in a carrier sentence. This is due to the effect of the adjacent voiceless sibilant sound [s] that initiates the word /sit/ 'six' following the two words /sabit/ and /ssabit/. [t] was found to be longer in /sabit/ than that in /ssabit/ when both words were said in isolation. However, the reverse was found in the event of pronouncing the two words in a carrier sentence. Vowel [i] in /sabit/ was almost of the same durational value as that in /ssabit/ when both words were spoken in isolation. The durational difference is noticeable when they were spoken in a carrier sentence, see tables 3.11 and 3.12.

Similarly, the durational measurements made for the two words /ħasan/ and /ħassan/ indicate that there is a considerable difference in the duration of the total word. In tables 3.13 - 3.16, it is shown that the word /ħasan/ has the mean duration of 636 msec when said in isolation and 487 msec when used in a carrier sentence; whereas the word /ħassan/ has the mean duration of 825 msec and 615 msec when it is spoken in isolation and in a carrier sentence, respectively. It is also observed that the duration of the geminate [ss] in /ħassan/ is twice that of the single [s] in /ħasan/ whether the two words are pronounced in isolation or in a carrier sentence. The voiceless pharyngeal fricative [ħ] is seen to be longer in the former word than that in the latter word when they are pronounced in isolation; a durational difference that is maintained when the two words are used in a carrier sentence, see tables 3.15 and 3.16. Besides, the nasal sound [n], occurring word-finally, has a considerably longer duration when the words /ħasan/ and /ħassan/ are spoken in isolation than when they are inserted in a carrier sentence. The duration of vowel [a] in the first syllable of the word /ħasan/

is almost the same as that of the word /ħassan/ when both words are enunciated in isolation or in a carrier sentence. The duration is found to be only 1 or 2 msec longer in /ħasan/ than that in /ħassan/. Similarly, the durational difference between vowel [a] in the second syllable of the two words is also found to be negligible; it is only 2 msec longer in /ħassan/ than that in /ħasan/, see tables 3.13 - 3.16. To our surprise, however, we find that vowel [a] in the second syllable is longer than that in the first syllable when both words are spoken in a carrier sentence. A fact that may be in complete contradiction with most researchers who claim that vowels occurring in stressed syllables are usually longer than those occurring in unstressed syllables. In this respect, one may suggest that presumably there is some effect on the duration of the preceding vowel [a] caused by the fact that [s] and [ss] are both voiceless, while [n] is voiced.[1] Researchers (e.g. Peterson and Lehiste, 1960; House, 1961; House and Fairbanks, 1953; Klatt, 1973) have long observed the fact that, for instance, in English vowels are generally longer before voiced consonants than before voiceless ones. Moreover, Chen (1970) presented further evidence when he found in his investigation on four languages (see section 2.4) that vowel duration varies as a function of the voicing of the following consonant and that a vowel is longer before a voiced consonant than that before a voiceless consonant.

In so far as the contrast between the words /darub/ versus /ddarub/ and /badal/ versus /baddal/ is concerned,

---

[1] Refer to what is said about /badal/ and /baddal/ later in this section.

table 3.17 - 3.20, for instance, indicate that the duration of the initial geminate consonant [dd] in the word /ddarub/ is considerably longer than its non-geminate counterpart [d] in the word /darub/ no matter whether these words are said in isolation or in a carrier sentence. The durational differences between the two consonants also made a substantial contribution to the overall difference of the duration of the two words. The total mean durations of the word /darub/, as shown in these tables, are 447 msec and 352 msec; whereas the word /ddarub/ has the total mean durations of 603 msec and 463 msec, respectively. The vowel sound [a] is seen to be slightly longer in /ddarub/ than that in /darub/ when both words are spoken in isolation; the difference is still negligible when the two words are spoken in a carrier sentence, see tables 3.19 and 3.20. Nevertheless, [a] is clearly longer in words pronounced in isolation than in those said in a carrier sentence. The flap [r] is also found to be slightly longer in /ddarub/ than that in /darub/ when the two words are spoken in isolation. It behaves conversely when they are spoken in a carrier sentence. Surprisingly, vowel [u], occurring in the second syllable, has exactly the same durational value in both words. It is noticeably longer in words pronounced in isolation than in those pronounced in a carrier sentence. Despite the negligible durational difference between word-final [b] in /darub/ and that in /ddarub/, it is found that it is longer in words uttered in isolation than in those spoken in a carrier sentence.

Tables 3.21 - 3.24 show us the durational measurements made for the words /badal/ and /baddal/. They also indicate that there is a considerable difference in the overall durations of the two words. The word /badal/ has the total mean duration of 497 msec when said in isolation and 385 msec when pronounced in a carrier sentence. The word /baddal/, on the other hand, has the total duration of 649 msec when spoken in isolation and 517 msec when used in a carrier sentence. Making a close inspection of these tables, one can see that the geminate [dd] is several times longer than its non-geminate partner [d] whether they occur in words spoken in isolation or in words pronounced within a carrier sentence. The word-initial [b] is longer in /baddal/ than in /badal/, and the durational difference is still clear between that pronounced in isolated words and another used in words spoken in a carrier sentence. On the other hand, the word-final lateral sound [l] shows exactly a similar trend to that of the word-final nasal sound [n] in /ħasan/ and /ħasan/, compare tables 3.21 - 3.24 with tables 3.13 - 3.16. By the same token, [l] is found to be considerably longer when enunciated in isolated words than in words used in a carrier sentence. The durational difference between vowel [a] in the first syllable of the two contrasting words is found to be negigible; [a] is only 3 msec longer before [dd] than that before [d] both in words spoken in isolation as well as in others spoken in a carrier sentence. Nevertheless, [a] is considerably longer in the first syllable than that in the second syllable in all cases.[1]

---

[1] Our discussion of the measurements of the sound components of the entire word, whether spoken in isolation or in phrasal context, leads us to the notion of temporal compensation which will be discussed in some detail in section 5.18.1.

Finally, it is worth mentioning that the durational measurements presented in these tables show clearly that consonants, whether they are singles or geminates, are longer when they are pronounced in words said in isolation than when they are uttered in words used in a carrier sentence. These tables also show us that the same consonants are longer when they occur word-initially than when they occur word-medially. The standard deviation, i.e. S.D., is found to be greater with the geminates than with the non-geminates. It is also greater with words spoken in isolation than with those spoken in a carrier sentence.

3.4 Discussion

Now, after making all these durational measurements referred to in the tables mentioned above, it seems quite evident that a comparison between single and geminate fricatives and/or plosives indicates that geminate consonants are considerably longer than their non-geminate counterparts no matter whether they are said in isolated words or in words produced in context. A geminate fricative, for instance, is almost twice as long as its single cognate when occurring in the same phonetic context; a fact that can be confirmed by having a glance at the durational measurements given in tables 3.1 - 3.4. In the case of the plosives, on the other hand, duration plays a distinct role in distinguishing single consonants from geminate ones. The closure duration of the geminate [dd] is about three or four times longer than that of the non-geminate [d]. The durational measurements given in tables 3.5 - 3.8 show this fact quite clearly.

It seems interesting to note that the differences in closure duration between geminate and non-geminate consonants found in this experiment for denti-alveolar plosives in I.C. Arabic are, in some way or another, comparable with those reported by Lehiste et al. (op.cit.) between single and geminate plosives in Estonian. According to them, differences between short and long medial geminates were firmly established. They were not only manifested by differences in duration but also "by emg duration and by differences in the number of emg peaks." [p.147]. They stated that closure duration served as a very effective parameter in distinguishing among the intervocalic consonant lengths in Estonian and in differentiating between pre- and post-junctural occurrences of consonants in English. However, the differences in I.C. Arabic for initial and medial contrasts between geminate and non-geminate denti-alveolar plosives are considerably greater than those found by Lehiste et al. (op.cit.) between geminate and non-geminate bilabial plosives in Estonian. They are even greater than those differences found by McKay (op.cit.) between single and geminate plosives in medial positions in Rembarrnga.

As for the contrasts between single and geminate fricative consonants, the results obtained in this experiment are in complete discrepancy with those reported by Obrecht (op.cit.) in the contrast between [ş] and [şş]. Obrecht states that the results of his experiment are less clear-cut in fricatives than in nasals and plosives. He believes that the reason is because of the nature of the phonetic opposition being manipulated. If his results are compared with those in our present experiment, we find that the case is the reverse. Measurements of duration for [s] and [ss] in both word-initial and word-medial positions indicate that the geminate denti-alveolar fricative [ss] is distinguished from the non-geminate [s] in that the former is characterized by a noticeable longer duration of noise. This fact, undoubtedly, emphasizes the significance of duration as being a cue factor in the contrast between single and geminate consonants. Besides, the mingograms obtained from the audio recordings of the stimulus items reveal typically that geminate fricatives are clearly characterized by a longer and more prominent burst of noise than the non-geminate ones, and by a longer and more abrupt closure for geminate plosives than their single counterparts. This phonetic observation was also reported by McKay (op.cit.) who stated that "these characteristics of the geminate stops may be considered indicators of fortis or tense articulation." [p.346].

In so far as the durations of the other sounds in the test words are concerned, it is found that in I.C. Arabic vowels occurring in stressed positions seem to maintain their original length whether they precede or follow a geminate consonant, and whether they exist in words spoken in isolation or in words pronounced in a contextual utterance. For instance, the vowel [a] has the mean duration of 99 msec in the word /sabit/ and 95 msec in the word /ssabit/ when both words are uttered in isolation. The same vowel has the mean duration of 81 msec in the two words when they are said in a carrier sentence. This fact also applies to vowel [a] when it precedes a geminate consonant or its single counterpart in all our test words whether they are spoken in isolation or in a carrier sentence, see tables 3.9 - 3.24.

Delattre (op.cit.) tried to find out whether geminate consonants would be preceded by shorter vowels than the corresponding single consonants. To his surprise, he found that "it was not the case, or at least not in a significant manner." [p.43]. He claimed that in French, German and Spanish, vowel duration results were comparable to those of English, i.e. vowels before geminates were on the average only slightly shorter than vowels before single consonants.[1] He added that a vowel preserved its original length despite a practical doubling of the following consonant's duration. Delattre, then, came to the conclusion that "in distinguishing a geminate from a single consonant, the duration of the preceding vowel is a negligible factor." [p.44]

In his investigation for medial plosive gemination, McKay (op.cit.) reached the same conclusion as that of Delattre. Although he found in some cases that shorter vowels occurred before the geminate plosives than before their single partners, he stated that "there was no structural linguistic evidence for phonemic vowel length distinction. The length of vowels has been interpreted here as insignificant." [p.346]. Looking at the measurements shown in tables 3.9 - 3.24 of our present experiment, it clearly seems that

---

[1] Underlining mine.

our findings are fully in agreement with those reported by Delattre and McKay. Our results indicate that vowels preceding single consonants are not essentially longer than those preceding geminate consonants, and that the duration of the preceding vowel can be considered as a negligible factor when discriminating between a single consonant and its geminate cognate.[1]

The final conclusion to be drawn from this experiment is, then, that duration is regarded as a major factor in discriminating geminate consonants from non-geminate ones in I.C. Arabic. Geminate consonants are shown to be considerably longer than their single partners no matter whether they occur word-initially or word-medially. Other secondary factors that help in the discrimination between the two phonetic categories may also exist. Proofs of such existence will be verified by further experiments. It is important to bear in mind the possible effects of producing words in isolation or in context, and our study of the duration of consonants in I.C. Arabic has been designed to clarify these effects.

---

[1] This point will be further dealt with in Chapter Six of this study.

Table (3.1) Durations (in msec) of word-initial [s] and [ss] used in words spoken in isolation.

Tokens	[s]	[ss]
1	171	318
2	186	295
3	194	326
4	217	349
5	194	295
6	186	341
7	194	331
8	171	326
9	182	341
10	171	326
Mean	187	325
S.D.	14	18

Table (3.2) Durations (in msec) of word-initial [s] and [ss] used in words spoken in a carrier sentence.

Tokens	[s]	[ss]
1	140	225
2	147	225
3	147	217
4	132	213
5	132	225
6	143	225
7	132	225
8	144	202
9	132	225
10	132	209
Mean	138	219
S.D.	7	9

Table (3.3) Durations (in msec) of word-medial [s] and [ss] used in words spoken in isolation.

Tokens	[s]	[ss]
1	116	248
2	116	248
3	128	248
4	124	287
5	116	248
6	120	295
7	124	225
8	128	240
9	120	248
10	120	248
Mean	121	254
S.D.	5	21

Table (3.4) Durations (in msec) of word-medial [s] and [ss] used in words spoken in a carrier sentence.

Tokens	[s]	[ss]
1	105	225
2	109	194
3	109	225
4	101	198
5	109	256
6	93	202
7	85	194
8	93	194
9	116	202
10	101	202
Mean	102	209
S.D.	10	20

Table (3.5) Durations (in msec) of word-initial [d] and [dd] used in words spoken in isolation.

Tokens	[d]	[dd]
1	85	217
2	62	225
3	78	256
4	85	236
5	70	209
6	62	186
7	85	225
8	78	221
9	54	225
10	74	235
Mean	73	224
S.D.	11	18

Table (3.6) Durations (in msec) of word-initial [d] and [dd] used in words spoken in a carrier sentence.

Tokens	[d]	[dd]
1	43	155
2	39	159
3	62	171
4	47	186
5	47	155
6	47	140
7	39	163
8	39	155
9	50	140
10	50	140
Mean	46	156
S.D.	7	15

Table (3.7) Durations (in msec) of word-medial [d] and [dd] used in words spoken in isolation.

Tokens	[d]	[dd]
1	31	155
2	31	147
3	31	147
4	27	147
5	23	143
6	23	163
7	31	155
8	23	155
9	31	163
10	25	155
Mean	29	153
S.D.	4	7

Table (3.8) Durations (in msec) of word-medial [d] and [dd] used in words spoken in carrier sentence.

Tokens	[d]	[dd]
1	16	132
2	23	140
3	31	132
4	27	140
5	23	140
6	16	155
7	39	147
8	16	143
9	23	155
10	23	132
Mean	24	142
S.D.	7	9

Table (3.9) Duration (in msec) of the word /sabit/ spoken in isolation.

Tokens	s	a	b	i	t	Total word
1	171	101	23	85	240	620
2	186	89	70	62	225	632
3	194	101	35	85	233	648
4	217	109	31	101	240	698
5	194	97	39	93	229	652
6	194	101	35	93	233	656
7	194	105	50	93	229	671
8	171	85	62	93	217	628
9	182	93	54	78	248	655
10	171	105	47	97	267	687
Mean	187	99	45	88	236	655
S.D.	14	8	15	11	14	25

Table (3.10) Duration (in msec) of the word /ssabit/ spoken in isolation.

Token	ss	a	b	i	t	Total word
1	318	85	54	70	217	744
2	295	93	54	85	194	721
3	326	93	39	78	225	761
4	349	105	39	85	163	741
5	295	85	47	81	225	733
6	341	85	47	89	171	733
7	333	85	54	89	225	786
8	326	93	47	89	236	791
9	341	109	54	101	202	706
10	326	120	47	101	236	830
Mean	325	95	48	87	209	755
S.D.	18	12	6	10	26	38

Table (3.11) Duration (in msec) of the word /sabit/ spoken in a carrier sentence.

Tokens	s	a	b	i	t	Total word
1	140	78	31	89	62	400
2	147	70	54	78	50	399
3	147	70	39	93	62	411
4	132	78	54	78	70	412
5	132	85	43	81	54	395
6	143	85	47	78	50	403
7	132	78	50	78	47	385
8	147	78	47	78	39	389
9	132	93	47	85	39	396
10	132	93	39	85	78	427
Mean	138	81	45	82	55	402
S.D.	7	8	7	5	13	12

Table (3.12) Duration in (msec) of the word /ssabit/ spoken in a carrier sentence.

Tokens	ss	a	b	i	t	Total word
1	225	78	39	81	47	470
2	225	70	47	78	78	498
3	217	85	39	62	62	465
4	225	78	47	78	54	482
5	233	89	31	85	62	500
6	228	85	39	70	70	492
7	225	89	39	78	78	509
8	209	78	39	81	54	461
9	225	81	47	78	47	478
10	213	81	47	78	85	504
Mean	223	81	41	77	64	486
S.D.	7	6	5	6	14	17

Table (3.13) Duration (in msec) of the word /ħasan/ spoken in isolation.

Token	ħ	a	s	a	n	Total word
1	147	85	116	58	233	639
2	155	85	116	54	213	623
3	132	78	128	47	209	594
4	136	93	124	62	240	655
5	151	81	116	54	217	619
6	163	85	120	62	194	624
7	147	78	124	47	240	636
8	171	78	128	50	217	644
9	171	70	120	62	240	663
10	155	78	128	54	248	663
Mean	153	81	122	55	225	636
S.D.	13	6	5	6	18	22

Table (3.14) Duration (in msec) of the word /ħassan/ spoken in isolation.

Token	ħ	a	ss	a	n	Total word
1	174	78	248	54	260	814
2	170	93	248	62	264	837
3	186	78	248	54	256	822
4	163	78	295	54	248	837
5	209	78	248	47	248	829
6	124	70	295	62	209	760
7	202	85	225	54	244	810
8	205	85	240	62	248	841
9	205	78	252	62	279	876
10	170	74	248	62	267	822
Mean	181	80	255	57	252	825
S.D.	26	7	23	5	19	29

Table (3.15) Duration (in msec) of the word /ħasan/ spoken in a carrier sentence.

Tokens	ħ	a	s	a	n	Total word
1	109	78	105	89	101	482
2	109	70	109	70	101	459
3	116	70	109	85	101	481
4	116	70	100	85	116	487
5	109	78	109	89	105	490
6	140	66	93	85	108	492
7	116	78	85	78	132	489
8	140	70	93	74	109	486
9	120	70	116	74	101	481
10	155	74	100	74	116	519
Mean	123	72	102	80	109	487
S.D.	17	4	10	7	10	15

Table (3.16) Duration (in msec) of the word /ħassan/ spoken in a carrier sentence.

Tokens	ħ	a	ss	a	n	Total word
1	178	62	225	85	85	635
2	136	62	202	93	89	582
3	155	66	225	85	93	624
4	132	85	198	85	101	601
5	182	70	256	93	85	686
6	109	85	202	70	116	582
7	143	62	194	62	101	562
8	155	78	194	70	93	590
9	182	62	202	85	109	640
10	163	70	205	93	116	647
Mean	154	70	210	82	99	615
S.D.	24	9	20	11	12	38

Table (3.17) Duration (in msec) of the word /darub/ spoken in isolation.

Tokens	d	a	r	u	b	Total word
1	85	124	62	66	132	469
2	62	124	66	78	140	470
3	78	116	54	66	120	434
4	85	112	70	78	128	473
5	70	109	74	50	120	423
6	62	120	70	50	124	426
7	85	116	62	58	132	453
8	78	116	78	50	132	454
9	54	112	70	50	128	414
10	74	124	70	62	124	454
Mean	73	117	68	61	128	447
S.D.	11	5	7	11	6	21

Table (3.18) Duration (in msec) of the word /ddarub/ spoken in isolation.

Tokens	dd	a	r	u	b	Total word
1	217	120	54	62	124	577
2	225	128	66	70	124	613
3	256	116	74	58	112	616
4	236	124	78	62	132	632
5	209	124	66	58	124	581
6	186	124	70	58	124	562
7	225	116	78	62	128	609
8	221	124	78	62	124	609
9	225	124	74	58	124	605
10	235	124	74	58	132	623
Mean	224	122	71	61	125	603
S.D.	18	4	8	4	6	22

Table (3.19) Duration (in msec) of the word /darub/ spoken in a carrier sentence.

Tokens	d	a	r	u	b	Total word
1	43	116	50	54	105	368
2	39	116	54	47	105	361
3	62	109	43	47	97	358
4	47	120	47	54	93	361
5	47	116	47	54	93	357
6	47	109	54	47	105	362
7	39	109	47	47	85	327
8	39	109	50	43	101	342
9	50	101	43	43	101	338
10	50	101	50	50	93	344
Mean	46	111	49	49	98	352
S.D.	7	6	4	4	7	13

Table (3.20) Duration (in msec) of the word /ddarub/ spoken in a carrier sentence.

Tokens	dd	a	r	u	b	Total word
1	155	116	47	47	101	466
2	159	116	47	54	97	473
3	171	116	47	47	78	459
4	186	109	47	54	105	501
5	155	116	50	47	101	469
6	140	109	47	47	85	428
7	163	109	47	43	93	455
8	155	112	47	47	89	450
9	140	109	47	50	105	451
10	140	109	54	54	116	473
Mean	156	112	48	49	97	463
S.D.	15	3	2	4	11	19

Table (3.21) Duration (in msec) of the word /badal/ spoken in isolation.

Tokens	b	a	d	a	l	Total word
1	85	109	31	58	202	485
2	93	124	31	47	202	497
3	105	124	31	66	217	543
4	74	124	27	58	217	500
5	105	101	23	78	209	516
6	89	109	23	62	194	477
7	93	101	31	62	159	446
8	89	101	23	62	225	500
9	85	93	31	70	233	512
10	58	101	35	58	240	492
Mean	88	109	29	62	210	497
S.D.	14	11	4	8	23	26

Table (3.22) Duration (in msec) of the word /baddal/ spoken in isolation.

Tokens	b	a	dd	a	l	Total word
1	93	120	155	62	202	632
2	89	128	147	70	202	636
3	109	116	147	78	240	690
4	93	124	147	70	264	698
5	147	97	143	54	182	623
6	101	109	163	78	209	660
7	93	93	155	74	217	632
8	70	116	155	62	248	651
9	85	109	163	70	221	648
10	39	109	155	62	259	624
Mean	92	112	153	68	224	649
S.D.	27	11	7	8	27	26

Table (3.23) Duration (in msec) of the word /badal/ spoken in a carrier sentence.

Tokens	b	a	d	a	l	Total word
1	54	101	16	120	109	400
2	58	101	23	89	101	372
3	50	112	31	93	109	395
4	58	105	27	101	105	396
5	54	101	23	85	109	372
6	54	101	16	97	101	369
7	62	121	39	97	89	408
8	47	116	16	89	97	365
9	47	116	23	85	101	372
10	58	124	23	89	105	399
Mean	54	110	24	95	103	385
S.D.	5	9	7	10	6	16

Table (3.24) Duration (in msec) of the word /baddal/ spoken in a carrier sentence.

Tokens	b	a	dd	a	l	Total word
1	81	101	132	97	78	489
2	70	116	140	101	101	528
3	62	109	132	97	101	501
4	58	109	140	112	78	497
5	78	124	140	101	105	548
6	70	124	155	116	93	558
7	62	109	147	109	81	508
8	54	112	143	109	89	507
9	69	112	155	109	93	538
10	70	109	132	101	85	497
Mean	67	113	142	105	90	517
S.D.	8	7	9	7	10	24

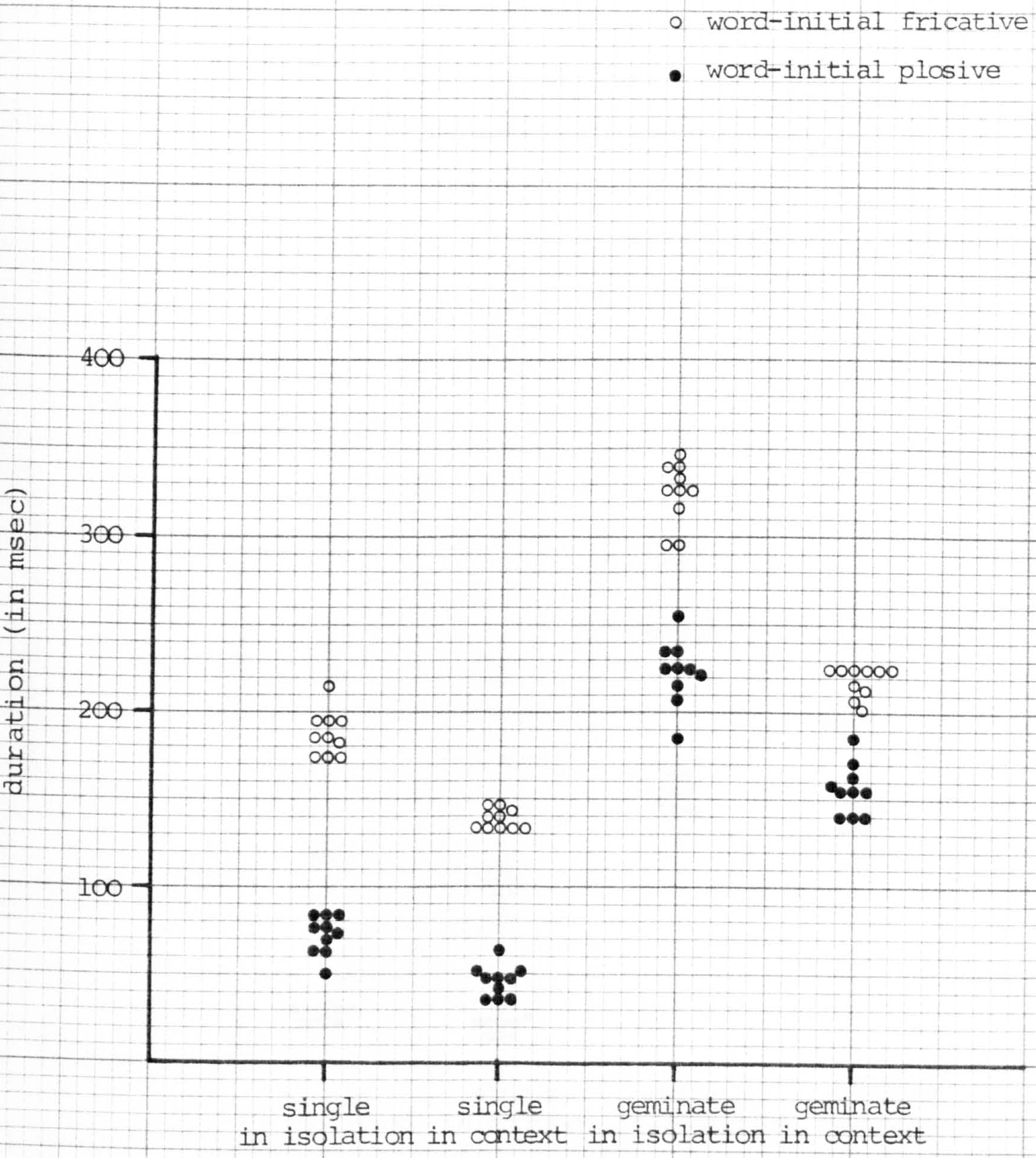


Fig. (3.6) Graph showing consonant duration in different contexts

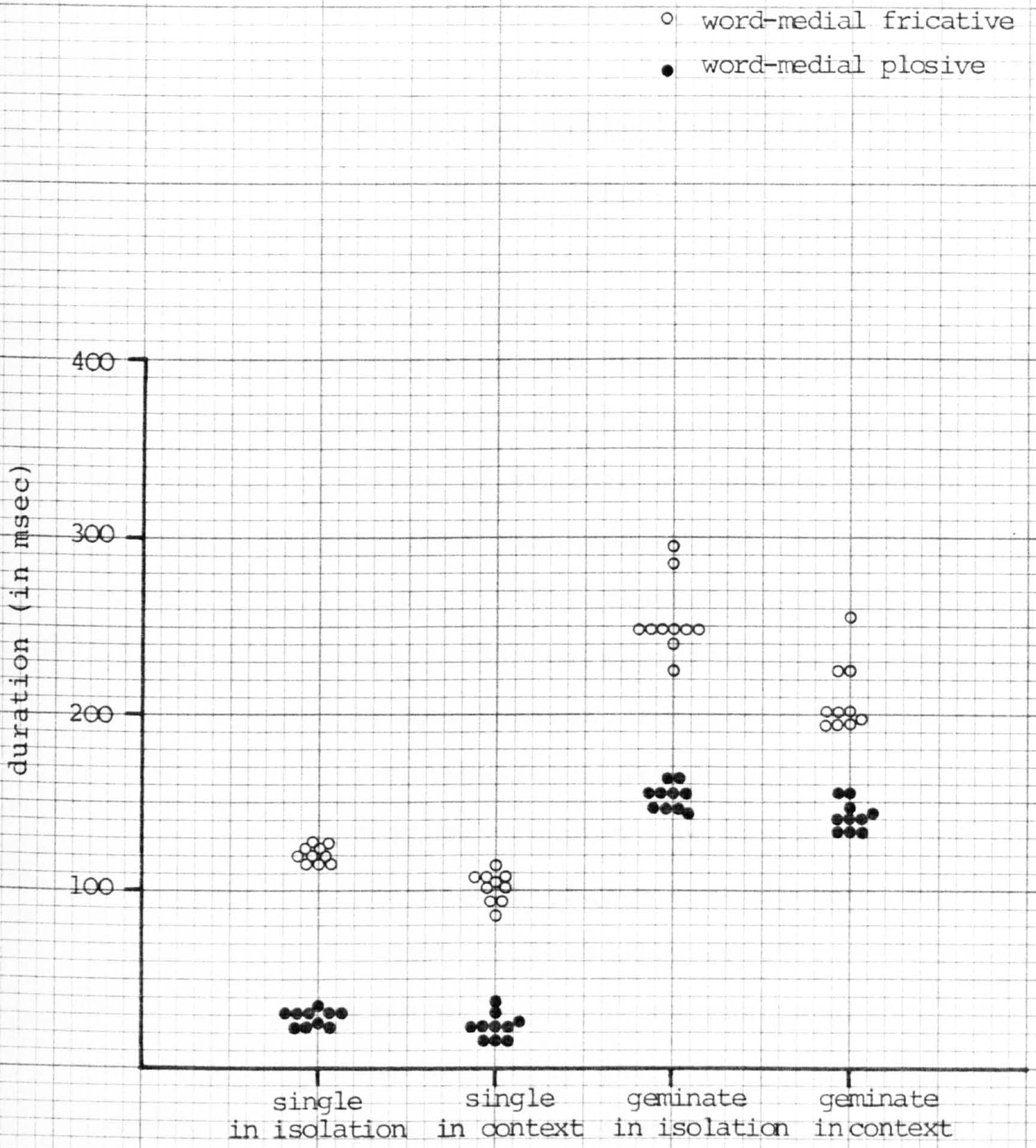


Fig. (3.7) Graph showing consonant duration in different contexts

CHAPTER FOUR

PERCEPTUAL STUDY OF GEMINATE CONSONANTS IN I.C. ARABIC

4.1 Introduction and Aims

The sound pattern of I.C. Arabic, as stated earlier, is characterized by the existence of distinctions in segmental durations of certain classes of speech sounds, viz. fricatives, plosives, affricates, nasals and approximants. The sounds of these classes are found in two different forms; either single or geminate. Like other languages, such as Japanese and Italian, contrasts between the single forms and the geminate forms are phonemic in I.C. Arabic.

In our previous chapter, it was shown that durational contrasts in I.C. Arabic play a significant role in the distinctions between various classes of speech sounds. Fry (1955), for instance, pointed out that segmental durations in some languages primarily carry both prosodic and emotional information, and that inter-segmental variations, if not insignificant, are almost always accompanied by more reliable and distinctive spectral cues.

It is of interest to mention here that there are some existing languages in which the segmental duration serves as the major cue for the distinction between certain classes of phonological units. In this respect, for example, the sound pattern of Japanese presents instances of considerable interest since all the vowels and some consonants possess longer counterparts that can be discriminated essentially by their durations. (Fujisaki et al., 1975).

It was the aim of the present chapter to deal with a synthetic investigation of the roles played by the segmental durations in the perception of cononants in I.C. Arabic. It was felt, at least at this stage, that doing some synthetic speech experiments would be of great importance to confirm the fact that the durational differences found in our acoustic experiment correspond to perceptual differences, and also to investigate phoneme boundaries between geminate and non-geminate consonants in I.C. Arabic.

4.2 Synthesis Along a Continuum

4.2.1 Investigation of Phoneme Boundaries:- A Critical Review

In some languages, such as Hindi, there exist two types of voiceless alveolar plosive consonants; one which is aspirated and the other which is unaspirated. To the listener of these languages there is apparently a division between the two types which is inoperative for the listeners of those languages in which the two types of sounds are non-existent. Fry (1964) states that "if we could produce a range of sound qualities filling, as it were, the space between aspirated and unaspirated [t] and send these one after the other to the Indian listener, we might expect that at some point in the series he would cease identifying the sound as aspirated and begin to identify it as unaspirated; he would, in other words, pass through a boundary between the two classes." [p.61].[1]

---

[1] Underlining mine.

In this regard, research groups at the Haskins Laboratories designed certain experiments to study this linguistic phenomenon for which they conveniently suggested the term 'phoneme boundary'. (Liberman et al., 1957). By this term they referred to "a boundary between classes and is not in any way related to the question of segmenting an utterance along the time dimension." (Fry, loc.cit.). For some time the Haskins Laboratories researchers have been trying to learn something about the perception of speech for which they have spent a great amount of their scientific work seeking for the acoustic cues on which that perception depends. Mainly, they have relied on techniques which enabled them to make controlled changes in various aspects of the acoustic pattern, and then to evaluate the effects of those changes on the sound as heard. They largely depended on the use of spectrograms for synthesizing and modifying the speech sounds. After having converted the spectrographic patterns into sounds, they listen to them to find out which of the experimental changes are important and which are not. The sound conversion obtainable from the spectrogram was always carried out by means of a machine which they call the 'Pattern-Playback' described with much detail in almost all their works on speech synthesis.

In accord with where and how the sounds are produced, the researchers at Haskins Laboratories divide the consonant cues into three classes. One class of consonant cues exists in sounds produced at the consonant constriction, such as the frictions of the fricatives and affricates as well as the bursts of the plosives. The constriction sounds are generally characterized by being produced "only during or just following the most nearly closed part of the consonant articulation." (Liberman, 1957, p.118). This is to say, the constriction sounds will quickly die away as their articulatory movement moves towards the next more open position of the following vowel. Pastore (1976) commented on Liberman and his co-workers' results of (1957) and stated that they favoured a learning explanation, but they did not reject completely the argument that "their results reflected the placement of phoneme boundaries at naturally occurring discontinuities in the discrimination functions." [p.253]. In order to examine the validity of this hypothesis, Liberman et al. (ibid.) suggested that their study should be replicated with non-speech stimuli which were, apart from that, identical to their synthetic speech stimuli. However, experiments carried out by Liberman et al. (1952) showed that the frequency position of the sound burst enables the listeners to discriminate among the voiceless plosives [p, t, k]. Harris (1954), on the other hand, stated the fact that the frequency location of the friction noise was an overwhelming cue for distinguishing [s] from [ʃ]. Other acoustic cues have also been found of some importance within this same group of constriction sounds such as duration and the nature of the onset of the noise which are considered as primary cues for various distinctions according to manner of articulation.

The second class of consonant cues contrast with the first one in that "the members of the second class result from the movement of the articulators." (Liberman, 1957). Besides, their excitation is generated in the larynx rather than at the point of consonant constriction. In an experiment conducted by Cooper et al. (1952) it was proved that the direction and extent of the second-formant transitions was an effective cue for distinguishing within the classes of plosive and nasal consonants. Later, O'Connor et al. (1957) found that the same variable, namely the direction and extent of the second-formant transitions, had a similar role with the approximants. Transitions of the third formants indicated, to some extent, identical effects to those of the second formants. Liberman (ibid.) reached the conclusion that "variations in direction and extent of second- and third-formant transitions are cues for the perception of various consonants according to place of production, while comparable variations of the first formant are cues for manner." [p.118]. In an earlier study, Delattre et al. (1955) suggested that for each consonant there are loci, viz. characteristic frequency positions, at which the formant transitions begin, or towards which they may be assumed to point. Liberman (ibid.) added that investigators of speech synthesis could successfully synthesize plosive consonants as well as nasal consonants only if they introduced a silent interval between the locus and the commencement of the transition.

The remaining third consonant cue results from the on-off action of a single fixed resonator in the nose. The fixed resonator and the corresponding acoustic cue "is an on-or-off nasal resonance that serves as an acoustic marker for the class of nasal consonants." (Liberman, op.cit., p.119). In a previous experiment, on the other hand, Liberman et al. (1954) had proposed that the nasal cue did not provide much of a basis for distinguishing the sounds within the class of nasal consonants, and the listener had to rely upon the transitions of both the second and third formants. In an important experiment accomplished by Miller and Nicely (1955) it was discovered that the cue elements and dimensions[1] were relatively independent of each other not only in the manner of their combination, but also in the way they were perceived.

In their first experiments, Liberman and his co-workers devised a plan to find out if there was any evidence indicating the existence of phoneme boundaries in listeners' behaviour. They were also interested in discovering the distinctions that were effective for the English listeners between the voiced plosives [b], [d] and [g]. Accordingly, they used their Pattern-Playback to generate speech-like sounds and to vary them in small gradual steps along an acoustic continuum comprising important cues for the perception of the voiced plosives. When the listeners were asked to label those sounds as [b], [d] or [g], they "normally tend by their responses to divide the continuum into three sharply defined phoneme categories, the shifts from one response, or phoneme label, to another being very abrupt."

---

1. By 'cue elements and dimensions' we mean the types of acoustic cues or the acoustic stimulus dimensions that have been found to be of some importance in the perception of the individual consonants of a language. For more details, see Liberman (1957).

(Liberman et al., 1957, p.359). It was the purpose of those experiments to determine the extent to which those sounds could be discriminated and also to see if the discrimination functions had sharp inflections that corresponded in position to the abrupt shifts, i.e. phoneme boundaries, in the labelling responses. Their experiments depended primarily on "setting up a series of test stimuli by means of which a speech synthesizer in which a periodic sound akin to the English vowel [a:] was initiated by a range of second formant transitions." (Fry, op.cit., p.62).

In earlier experiments Liberman and his colleagues at the Haskins Laboratories had shown that a minus second formant transition associated with the same vowel, viz. [a:], was perceived by most English listeners as the syllable /ba:/; whereas the absence of the transition or a slight plus transition was heard as the syllable /da:/, and a more pronounced plus transition was identified as the syllable /ga:/. In synthesizing these syllables "it was possible to proceed by very small steps through a wide range from a large minus second formant transition through the zero transition to a large plus transition." (Fry, loc.cit.). The resulting synthesized syllables were, then, set in random order and presented to English listeners who were instructed to state whether they heard [b], [d] or [g]. The reaction of the listeners' labelling over certain parts of the continuum took the form of regular patterns; that is to say, they were clearly identified as [b], [d] or [g]. Still, there were uncertain parts of the continuum where the stimuli were labelled almost haphazardly. These regions of uncertainty constituted what Liberman and his colleagues called as the 'phoneme boundary'. Lane (1965) states that in employing synthetic vowel stimuli Liberman et al. (1963) found, on an identification test, reasonably well-defined and clear boundaries between phoneme categories, and on a discrimination test, peaks at the phoneme boundaries, i.e. discrimination between adjacent stimuli in the vowel continuum, are better when the stimuli lie near phoneme boundaries.[1]

When we come to speak about phoneme boundaries in relation to duration, we can refer to those experiments which were principally designed to investigate them as well as to those in which the notion of the phoneme boundary was implied. For example, in the experiments conducted by Denes (1955), it was shown that the duration ratio of the periodic and the noise segments of a syllable was considered as a cue for distinguishing final [s] and [z] in English, and a clearly recognizable phoneme boundary, where the ratio of noise to periodic sound was 1.0, was easily captured by the listeners' auditory judgments in those experiments.

Lisker (1957), on the other hand, examined the significance of duration for the distinction between the two English plosives [p] and [b] occurring in intervocalic positions. The results of his investigation indicated that the closure duration of [p] and [b] in the two words 'rupee' and 'ruby' was a significant cue for his subjects to distinguish the two words in the listening tests. He found that the duration of the

---

[1] Lane's criticism of the findings of the Haskins Laboratories researchers can be referred to in section 4.6.1 of this study.

silent interval between the first and the second syllables in each of the two words was a sufficient acoustic cue, and discovered that decreasing the silent interval would cause the listener to hear the word as 'rabid'; whereas increasing it would cause the perception to change into 'rapid'. Lisker came to the final conclusion that "for the particular isolated words that he used there was a phoneme boundary with respect to duration of closure in the region between 75 msec and 130 msec, that is to say, all closures shorter than 75 msec were heard as /b/ and all closures longer than 130 msec were heard as /p/." (Fry, 1964, p.66). Liberman (1961) supported Lisker's conclusion. He obtained data for labelling and discriminating the two English words 'rapid' and 'rabid'. He found that the phoneme boundary which was apparent in the labelling data was matched by a maximum in the ability to discriminate. Fry (op.cit.) suggested that the distinctness of a phoneme boundary might depend in part on the degree of articulatory difference between one phoneme and another. He added that "this factor undoubtedly has some influence on perception and will be quite strong in the case of most consonants, at least." [p.67].

In other experiments dealing with synthetic plosive consonants (e.g. Liberman et al., 1957; Liberman et al., 1961; Bastian et al., 1961), it was found that discrimination of a constant acoustic difference was considerably better across phoneme boundaries, between [b] and [d] and between [d] and [g], than in the middle of the phoneme categories. It was reported by Liberman et al. (1961) that "the discrimination was, in fact, relatively so good across the boundaries and so poor within the categories as to suggest that the subjects could only respond to these sounds categorically (i.e. as phonemes) and could hear no other differences among them." [p.379]. To get the appropriate measurements, the researchers had first to identify acoustic variables which were sufficient cues for the perceived phonemic distinctions. After selecting an appropriate acoustic cue for each phonemic distinction, they planned a series of synthetic patterns in which that acoustic cue was altered along a single continuum through a large range as to include those phonemes under investigstion. Aiming at measuring discriminability, they organised the patterns into what is called ABX procedure, and the listeners were instructed to tell whether X was identical with A or with B; A and B were always different and X was always identical with either A or B. In order to establish the phoneme boundary, the researchers decided to present the patterns with instructions to identify each one as [b], [d] or [g] in the first experiment, as [d] or [t] in the second, and as [sl] or [spl] in the third. They also intended to present the synthetic speech stimuli arranged in the form of ABX triads to measure percent correct discrimination within pairs of stimuli extracted from the speech continuum. Lane (op.cit.) commented on this experimental procedure saying that the researchers expected to find that two stimuli extracted from the same phoneme category were less often correctly discriminated than two stimuli extracted from opposite sides of the phoneme boundary.

Cooper et al. (1952) contended that a careful inspection of actual spectrograms suggested that one of the variables that might enable a listener to differentiate [p], [t] and [k] was "the position along the frequency scale of the brief burst noise which constituted the acoustic counterpart of the articulatory explosion." [p.597]. In an attempt to isolate this variable and to determine its role in perception, Cooper and his co-workers designed a series of schematized burst-plus-vowel patterns in which bursts at each of 12 frequency positions were arranged in pairs with any of the seven vowels [i, e, ɛ, a, ɔ , o, u]. The bursts were kept constant as to size and shape, and the vowels were composed of merely two formants.[1] The combinations of burst and vowel were transferred into sound and, then, randomized and presented to a number of English listeners who were asked to identify the initial component of the syllable as [p], [t] or [k].

The results obtained from this experiment showed that the listeners' identification varied in accord with the frequency position of the burst. The variable of the frequency position of the burst provided the listeners with a basis for distinguishing among [p], [t] and [k]. The results also indicated that high frequency bursts were heard as [t] for all seven vowels; whereas, bursts at lower frequencies were heard as [k] when they were on a level with or slightly above the second formant of the vowel; otherwise they were heard as [p]. The investigators then concluded that "the transitions of the first formant appear to contribute to voicing of the

---

[1] For further information, see Delattre et al. (1951), pp.30-36.

stop consonants, while transitions of the second formant provide a basis for distinguishing among /b/, /d/ and /g/ or their cognates /p/, /t/ and /k/." [p.600].

In a remarkable endeavour to synthesize speech from discrete segments, Peterson et al. (1958) used what they termed "dyads'[1] to produce speech rather than spelling; therefore, including units which exceeded one phoneme in length. Their experiments on consonant patterns, such as plosives, showed that a primary cue for the perception of these and certain other consonants was the relatively rapid shift in the formant frequencies. These shifts, i.e. transitions, are among the most important cues to the perception of many of the consonants. Liberman et al. (op.cit.) emphasized the fact that the perceptual function of the transition is to provide important and sometimes essential information for phoneme identification.

Now after all that has been reported about the studies on speech synthesis by the researchers of the Haskins Laboratories, one can finally agree with what Fry (1973) has suggested in saying that their experiments have shown how in judging English plosives, listeners may make use of the frequency of a burst of noise, the extent and direction of a formant transition, the time of onset of a formant, and the duration of a particular part of the sound-wave.

---

1. A 'dyad' is defined as a segment which contains parts of two phones with their mutual influence in the middle of the segment. (Liberman et al., 1959, p.1491).

4.3 Experimental Procedure

4.3.1 Description of the Program

The experiments on speech synthesis which will be described in this chapter depend on a system which has been developed in the Phonetics Laboratory at Leeds University by Peter Roach in collaboration with Janak Gandhi. The general principles and the detailed operating instructions of how to use the program, namely the SPS256, to synthesize speech are explained in an instruction manual prepared by the two researchers.[1] Gandhi and Roach (1980) contend that in spite of the fact that speech synthesis is a slow and laborious job, the program has been designed to let the investigator "gain experience in manipulating acoustic phonetic parameters as simply as possible." [p.1].

The computalker speech synthesizer used in our present experiments is controlled by nine parameters; four of which are 'amplitude' parameters which are related to the four basic types of sound that can be generated by the speech synthesizer. The four amplitude parameters are:

AV	Voicing Amplitude
AH	Aspiration Amplitude
AF	Frication Amplitude
AN	Nasal Amplitude

The two parameters AV and AN are related to voicing; while parameters AF and AH are related to noise. Any value for any of these parameters must be within a limited range. To get

---

[1] For full details about the SPS256 program and the various functions of the commands, see appendix 1.

a vowel sound, for example, the AV amplitude parameter value has to be high enough to produce audible vowel-type sound and the other three amplitude parameters must be at zero. (Gandhi and Roach, op.cit.). On the other hand, to get a voiceless fricative other than [h] the AF amplitude parameter value has to be above zero and the other amplitudes at zero.

The last five of the nine synthesizer control parameters have a frequency-related control function. These are:

FO	Frequency of Voicing
F1	Frequency of Formant One
F2	Frequency of Formant Two
F3	Frequency of Formant Three
FF	Frequency of Frication

Of these, Frequency of Voicing parameter, viz. FO, moves as a direct exponential function of the control value, and it is a characteristic of all voiced sounds. The other four parameters, i.e. F1, F2, F3 and FF, move as an inverse exponential function of their control values, and they are related to spectral characteristics. They are 'formant' parameters. The first three are necessary for producing vowels and vowel-like sounds; whereas the fourth one is necessary to produce differences in fricative quality.

For the purpose of producing reasonably clear speech, the speech synthesizer must be fed with values for each parameter at least 100 times per second. In this respect, Roach (1980) states that "for each of these parameters a byte has to be sent to the appropriate part at regular intervals. The usual update rate is 10 msec, though for the highest quality 5 msec is preferable. A 10 msec frame rate requires data output at 900 bytes per second, and it would obviously be too laborious to input all these values manually." [p.11]. It is the job of the computer to produce these enormous quantities of numbers, and it is the job of the experimenter to feed the computer with the information it needs so as to generate the correct numbers. In fact, the most important thing is to load frequency and amplitude values for the points in time at which spectral changes have to be brought about and let the program generate the data frames for the intervening time intervals. The high output rate and the large number of data bytes to be stored makes some machine code programming essential. (Roach, op.cit.). For instance, if the experimenter wishes to produce a steady vowel all he has to do is to feed the computer with the correct values out of the nine parameters, and instruct it how long to keep sending these values to the speech synthesizer every hundredth of a second. A tenth parameter value, namely (PD), specifying the duration must also be added to the other values of the nine parameters.

In speech synthesis, a lot of practical experience is of ultimate necessity so as to produce a satisfactory synthetic output. However, there is a basic principle that the experimeter should pay attention to. More specifically, it is the experimenter who decides the points in time at which important changes in acoustic structure take place, and whether there is a gradual or an abrupt change from the values of one point to the values of the following point. (Gandhi and Roach, op.cit.). The number of points used depends, to a certain extent, upon the accuracy of specifying the acoustic structure of the speech to be used. On average, three or four points per phoneme are necessary. Fig. 4.1 is a schematic diagram of the synthesizer employed in the present experiments on speech synthesis.

4.3.2 Operating Instructions and Program Commands

In speech synthesis the role of the computer is primarily to read parameter values at discrete points, calculate values between points and directly transmit them to the synthesizer at a specified rate. In addition to all these functions, the SPS256 program also provides facilities for editing, displaying and saving data. It handles all the functions required before and after synthesis. The key functions required of the program, as stated by Roach (1980), are:

(a) To generate control parameter values for data frames and later output them to the part of the synthesizer.

(b) To control the output rate of data frames.

Basically, the program is a monitor. It accepts the experimenter's commands from the computer keyboard, carries out the required job and waits for the next command. All the commands are in the form of single letters and the parameter names are in the form of two characters, see section 4.3.1. It sometimes happens that the program refuses to accept certain commands and the data values. In such cases the

Fig. (4.1) **Block diagram of the synthesizer used in this experiment**

Fig. (4.1) A block diagram of the synthesizer. It shows the organization of the synthesizer which includes a wide variety of parameters. The voicing frequency and amplitude are set by parameters FV and AV. The noise pulses of stop consonants are generated with the programmable gain element AH. The fricative resonator with amplitude AF and frequency resonance FF are used to generate fricatives like "s" and "sh". The normal vowel sounds are generated by control of the formant frequencies F1, F2 and F3, and a nasal resonator with amplitude AN and fixed frequency characteristics is used to add varying amounts of nasal sounds. The result of signals processed through the nasal, formant and fricative paths is summed by a final operational amplifier and used to drive the output speaker.

computer displays a message indicating the reason for the refusal on the screen and the program waits for the next command or for a new data value. For example, when the computer screen displays the sign '*', it means that the program is ready to accept one of the commands, and when it displays the sign '?' it denotes that the experimenter is being instructed to input a data value or a parameter name as specified by whatever precedes '?' on the screen. (Gandhi and Roach, op.cit.).

The operating commands which the SPS256 program accepts are twelve in number. Briefly, they are stated below with their basic functions.[1]

Commands	Functions
C	Create the intermediate data frames.
D	Display of points data on screen.
E	Exit from the program.
G	Graph parameter on printer.
I	Input of data to computer.
L	List data for all points.
M	Modify one parameter value for one point at a time.
O	Overall modification.
P	Play the synthesized speech through the Ferrograph-Logic 7.
R	Record the synthesized speech on the Ferrograph-Logic 7.
S	Save points data in digital form on tape.
V	VDU graph.

---

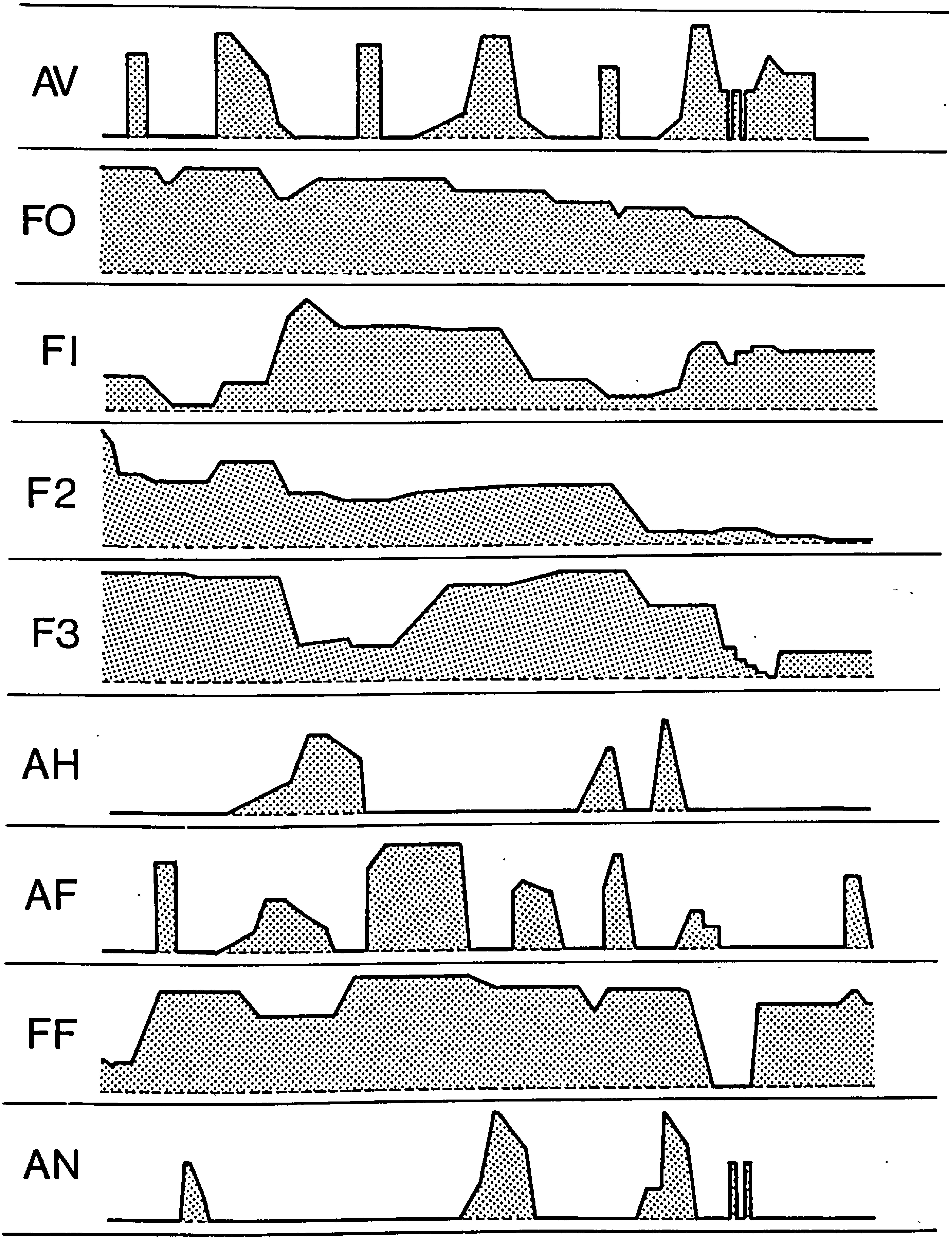
[1] For detailed information about the description of these commands, see appendix 1.

4.3.3 Synthesis Parameter Data Used

In order to produce the synthesized speech for the two utterances /'kitbi 'hassan 'sit mar'raat/ and /'kitbi 'baddal 'sit mar'raat/ only 43 points were decided for the first utterance and 41 points for the second utterance on spectrograms of real speech. For each point the appropriate values of the nine parameters were estimated after making the due measurements on the spectrograms. A special coding form was prepared into which these parameter values were written down. Setting the synthesis program into operation, the nine parameter values for each point were fed into the computer, which in turn would calculate the intermediate data frame values generated. Whenever we felt that the values of the nine parameters need to be inspected to see whether they were exactly within their appropriate range, the computer was asked to do so by typing in the command D, i.e. display of data on screen. Accepting this command, the computer would immediately show the nine parameter values for every five points on its screen. By typing the command V, i.e. graphic display on VDU, a visual representation of certain parameters at some chosen segments would also be obtainable on the screen or by the printer to check the accuracy of the range for the entered values. (See Fig.4.2).

For the purpose of creating our two listening tests for the two experiments to investigate phoneme boundaries between [s] and [ss] on the one side, and between [d] and [dd] on the other, the parameter value of merely one point, namely PD, was intentionally modified; while the values of the other

Fig. (4.2) Schematic Diagram of Synthesizer Parameter Values for One Version of /'kitbi 'ħassan 'sit mar'raat/.


AV
FO
F1
F2
F3
AH
AF
FF
AN

points remained unchanged. The modifications took place at points No. 15 and No. 14 in the first and second experiments, respectively. (See Figs. 4.3 and 4.4).

At certain stages in our two experiments, and after listening to the synthesized speech via the Ferrograph-Logic 7 for several times, spectrograms for the synthetic output were made and then compared to those of real speech. Modifications were made for some parameter values at those points which seemed less identical, if compared, to those on real speech spectrograms when we felt that the synthetic output was perceptually unsatisfactory. The modifications were continued until a satisfactory version of the synthetic speech was heard. Once we felt that the overall outcome of the synthesized speech was reasonably intelligible, we began creating the data for our listening tests. However, although the similarity between real and synthetic speech was highly acceptable, some of the acoustic differences are immediately visible on spectrograms.

If one scrutinizes the spectrographic representations for the two carrier sentences /'kitbi 'ħasan 'sit mar'raat/ and /'kitbi 'badal 'sit mar'raat/ as pronounced by a real speaker and those produced by the speech synthesizer, see Figs. 4.5-4.8, one can judge the acceptability of the products generated by speech synthesis. In this regard, Roach (1980) remarks that despite the fact that the system has only recently been brought into use, yet he is confident that he has a system that will be very useful. He also adds that "the quality, as judged by several listeners experienced in listen-

PT.	AV	F0	F1	F2	F3	AH	AF	FF	AN	FD
001	000	110	360	2750	2750	000	150	2750	000	015
002	001-	110	360	2700	2750	175	000	2700	000	030
003	150-	110	360	2100	2750	000	000	2750	000	040
004	000	110	360	2100	2750	000	001-	5000	000	060
005	000	110	175	2000	2720	000	150-	5000	001-	030
006	001-	105	175	2000	2720	000	000	5000	150	070
007	180	110	175	2000	2720	000	000	5000	050	020
008	180	110	320	2240	2720	000	000	5000	000	115
009	100	110	320	2240	2720	050	025	4000	000	025
010	020	100	800	1680	2560	100	070	4000	000	040
011	000	100	900	1720	2176	100	070	4000	000	095
012	000	105	750	1600	2200	075	020	4000	000	010
013	180-	105	750	1550	2100	000	001-	4000	000	055
014	000	105	750	1550	2100	000	175	6000	000	030
315	000	105	750	1550	2100	000	200	6000	000	180 ←
016	020	105	725	1640	2760	000	195	6000	000	025
017	180	105	725	1640	2760	000	000	6000	000	070
018	180	100	725	1640	2760	000	000	5520	100	020
019	025	100	725	1640	2760	000	001-	5520	250	085
020	000	100	368	1680	2800	000	100	5520	160	020
021	000	100	368	1680	2800	000	120	5520	000	095
022	001-	100	368	1680	2800	100	100	5520	000	020
023	120-	95	368	1680	2800	000	000	5000	000	040
024	000	95	368	1680	2800	000	001-	5500	000	080
025	000	95	240	960	2500	000	100	5500	000	015
026	000	95	240	960	2500	050	150	5500	000	020
027	000	90	240	960	2500	125	000	5500	050-	020
028	000	93	240	960	2500	000	000	5500	250	060
029	030	93	250	960	2500	000	000	5500	200	020
030	200	93	440	960	2500	000	000	5500	000	040
031	200	93	499	960	2500	000	000	2000	000	030
032	100-	93	500	960	1900	000	040-	2000	001-	025
033	001-	90	400	990	1900	000	001-	2000	100-	005
034	100-	90	400	990	1850	000	020-	2000	001-	025
035	001-	90	400	990	1840	000	001-	2000	100-	005
036	100	90	450	990	1790	000	000	2000	000	020
037	100	90	450	990	1775	000	000	2000	000	020
038	150	90	490	920	1900	000	000	5250	000	060
039	120-	84	480	910	1900	000	000	5250	000	080
040	000	75	480	910	1900	000	001-	5250	000	085
041	000	75	480	910	1900	000	100	5250	000	015
042	000	75	480	910	1900	000	100	5350	000	020
043	001-	75-	480-	910-	1900-	001-	001-	5250-	001-	000

" *kitbi kassan sit marraat* "

Fig. (4.3) Computalker Parameter Data

PT.	AV	F0	F1	F2	F3	AH	AF	FF	AN	FD
001	000	110	360	2750	2750	000	150	2750	000	015
002	001-	110	360	2700	2750	175	000	2750	000	030
003	150-	110	360	2100	2750	000	000	2750	000	040
004	000	110	360	2100	2750	000	001-	5000	000	060
005	000	110	175	2000	2720	000	150-	5000	001-	030
006	001-	105	175	2000	2720	000	000	5000	150	070
007	180	110	175	2000	2720	000	000	5000	050	020
008	180	110	320	2240	2720	000	000	5000	000	115
009	100	110	320	2240	2720	000	000	5000	050-	015
010	020-	105	175	2000	2720	000	000	5000	255	075
011	000	105	650	1600	2720	000	000	5000	050	020
012	180-	105	650	1760	2720	000	000	5000	000	050
013	100	110	650	1760	2720	000	000	5000	050-	025
014	020-	110	180	1760	2720	000	000	5000	255	100 ←
015	000	110	480	1920	2800	000	000	4000	050	020
016	180-	105	640	1760	2640	000	000	4000	000	080
017	100	100	320	1600	2600	000	000	4000	100	095
018	000	100	368	1680	2800	000	100	5520	050	020
019	000	100	368	1680	2800	000	120	5520	000	095
020	001-	100	368	1680	2800	100	100	5520	000	020
021	120-	95	368	1680	2800	000	000	5000	000	040
022	000	95	368	1680	2800	000	001-	5500	000	080
023	000	95	240	960	2500	000	100	5500	000	015
024	000	95	240	960	2500	050	150	5500	000	020
025	000	90	240	960	2500	125	000	5500	050-	020
026	000	93	240	960	2500	000	000	5500	250	060
027	030	93	250	960	2500	000	000	5500	200	020
028	200	93	440	960	2500	000	000	5500	000	040
029	200	93	499	960	2500	000	000	2000	000	030
030	100-	93	500	960	1900	000	040-	2000	001-	025
031	001-	90	400	990	1900	000	001-	2000	100-	005
032	100-	90	400	990	1850	000	020-	2000	001-	025
033	001-	90	400	990	1840	000	001-	2000	100-	005
034	100	90	450	990	1790	000	000	2000	000	020
035	100	90	450	990	1775	000	000	2000	000	020
036	150	90	490	320	1900	000	000	5250	000	060
037	120-	84	480	910	1900	000	000	5250	000	080
038	000	75	480	810	1900	000	001-	5250	000	085
039	000	75	480	910	1900	000	100	5250	000	015
040	000	75	480	910	1900	000	100	5350	000	020
041	001-	75-	480-	910-	1900-	001-	001-	5250-	001-	000

" *kitbi baddal sit marraat* "

Fig. (4.4) Computalker Parameter Data

ing to synthetic speech, was quite good." [p.11].

Depending on the results we have obtained from our two experiments on speech synthesis, we entirely agree with Roach's statement that 'the quality was quite good.' This statement was shown to be true when we presented our synthetic speech to a number of native speakers of I.C. Arabic who served as subjects in the two listening tests and whose responses were judged with the help of a computer. (See section 4.5). Nevertheless, an inspection of the spectrograms representing the synthesized speech shows that F3, in general, has a very weak energy or it even vanishes. From the perceptual point of view, this acoustic phenomenon does not greatly affect the subjects' capability of perceiving the synthetic utterances as being reasonably identifiable.[1] For instance, the two fricative consonants [s] and [ss] in /ħasan/ and /ħassan/ show a noticeable noise energy along the frequency scale, and the greatest energy of noise gathers around 6000 - 7000 Hz. (See Fig. 4.9). In so far as the two plosives [d] and [dd] in the words /badal/ and /baddal/ are concerned, the spectrograms show no indication of any sort of voicing energy on the frequency scale. However, this fact does not greatly affect the acceptability of [d] and [dd] as being recognized as voiced plosives in the listening tests as well. The spectrograms expose the same feature, i.e. the absence of voicing energy, as to the voiced bilabial plosive [b] occurring word-initially in /badal/ and /baddal/, and word-medially in /kitbi/. Besides, what differentiates the voice-

---

[1] Some of the subjects even declared to the present author that they could hardly find any difference between a natural speech and what they had heard.

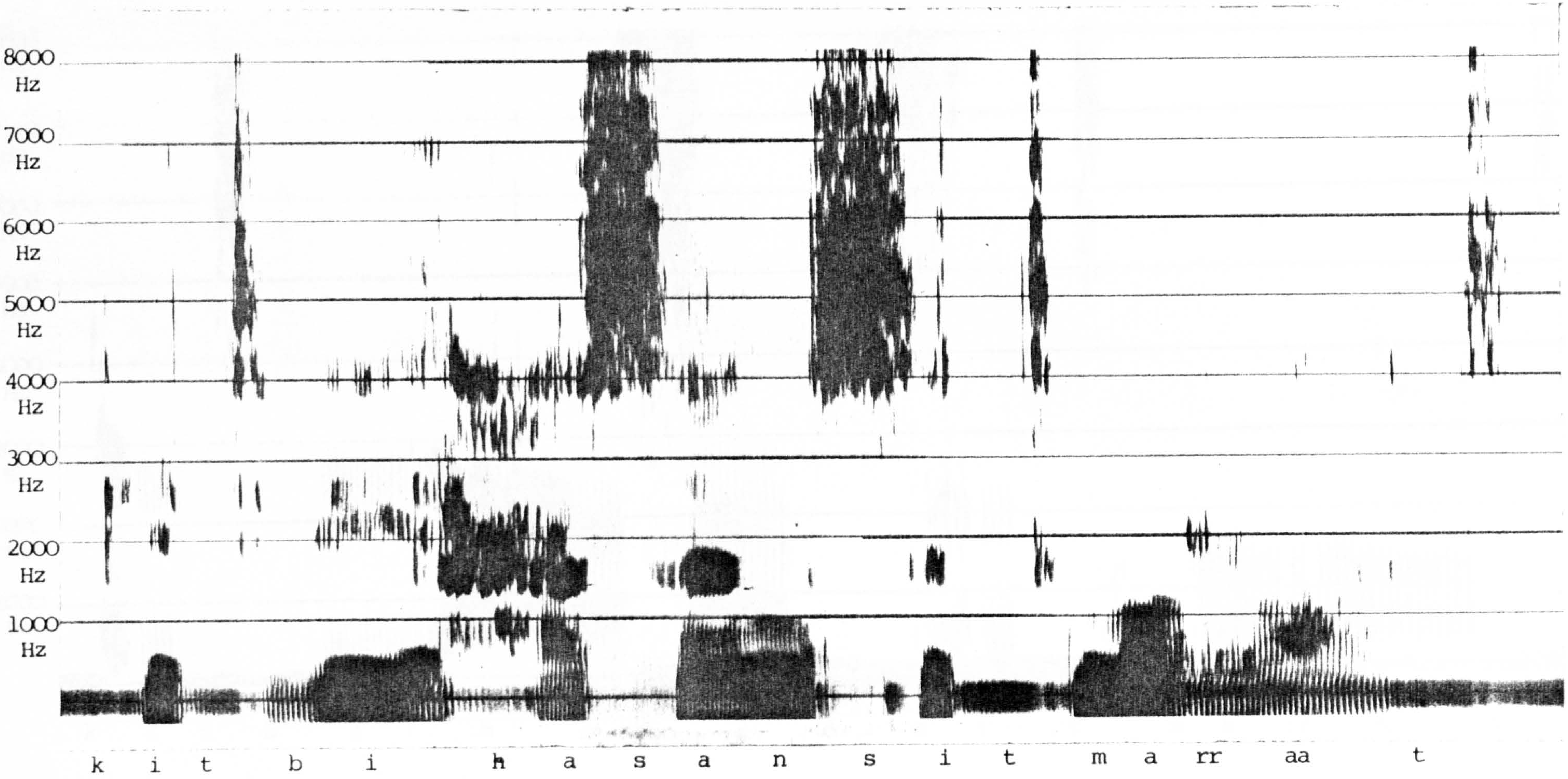


Fig. (4.5) Spectrogram for /kitbi ħasan sit marraat/ as said by a real speaker

Fig. (4.6) Spectrogram for /kitbi ħasan sit marraat/ as produced by a terminal-analog speech synthesizer

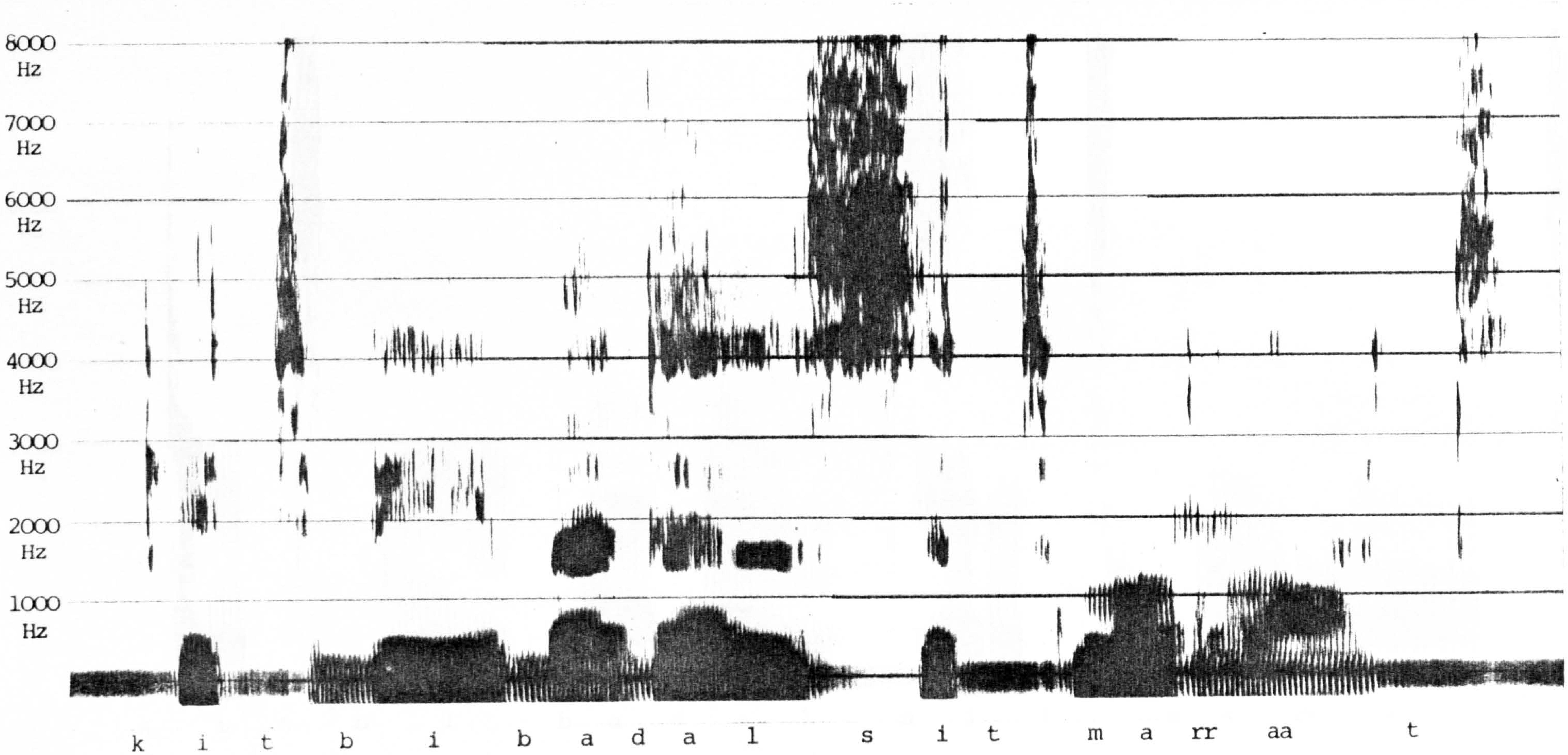


Fig. (4.7) Spectrogram for /kitbi badal sit marraat/ as said by a real speaker

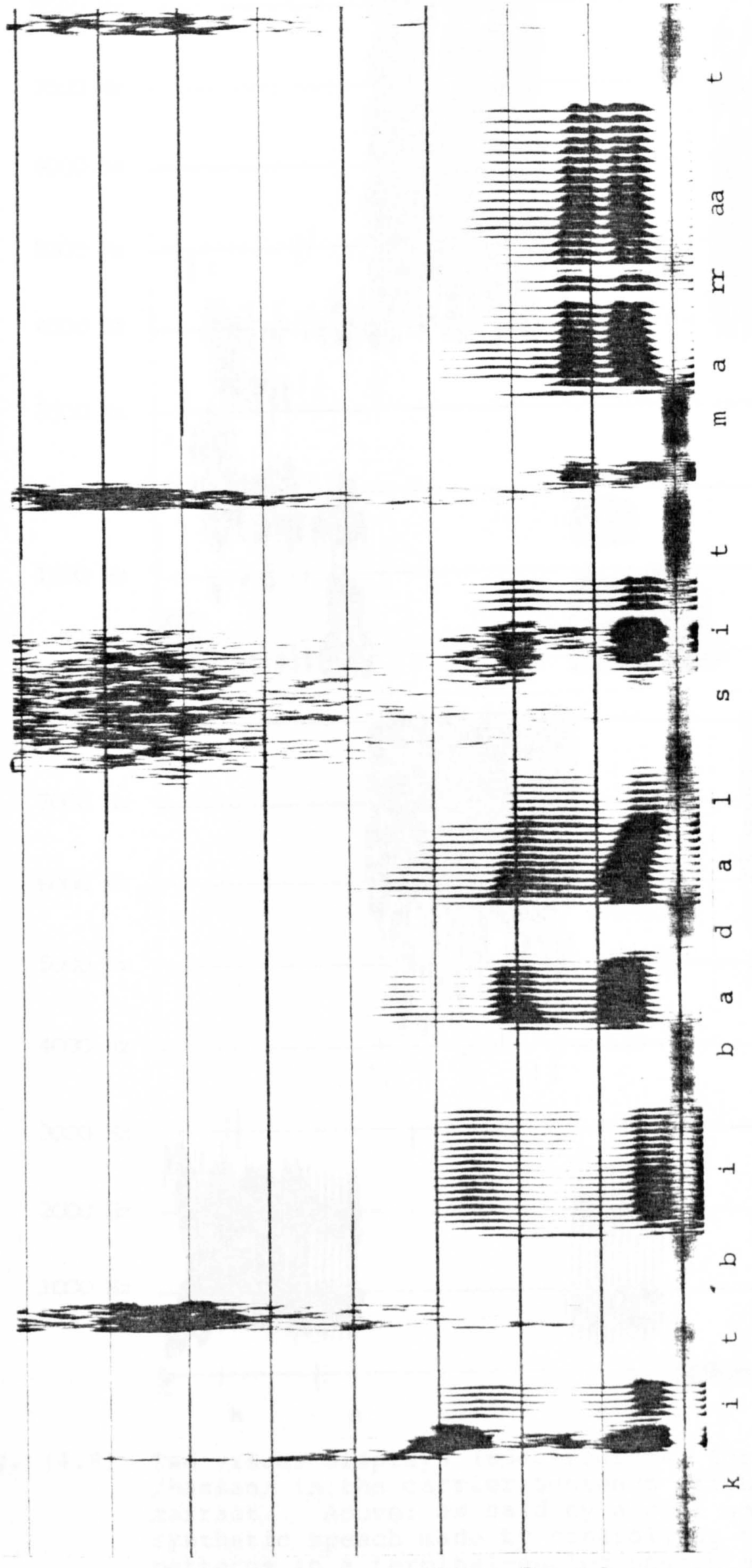


Fig. (4.8) Spectrogram for /kitbi badal sit marraat/ as produced by a terminal-analog speech synthesizer

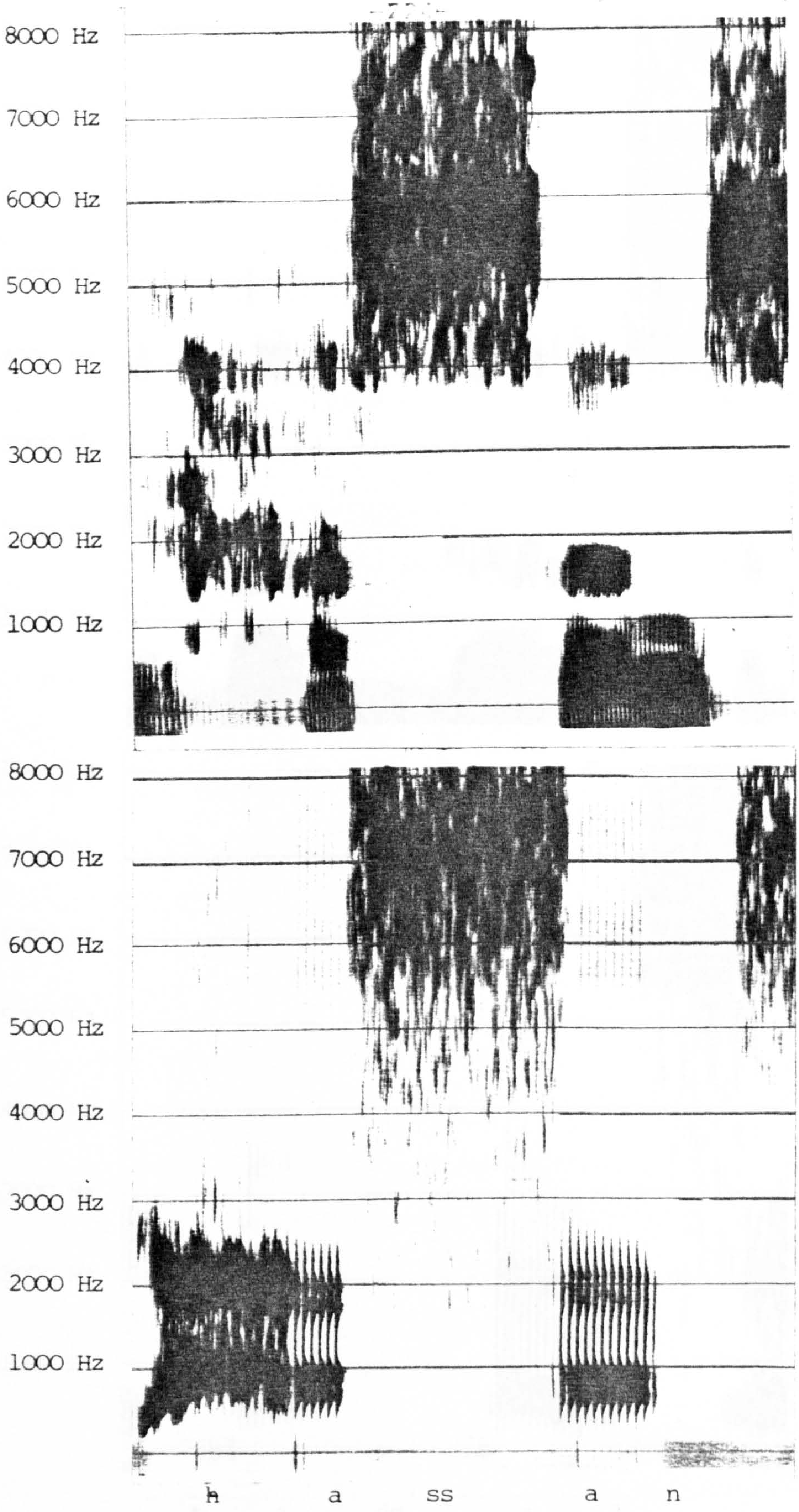


Fig. (4.9) Two visual displays (spectrograms) for the word /ħassan/ in the carrier sentence /kitbi ħassan sit marraat/. Above: as said by a real speaker; below: synthetic speech made by controlling the sound patterns in a terminal-analog speech synthesizer driven by the Department's PET computer

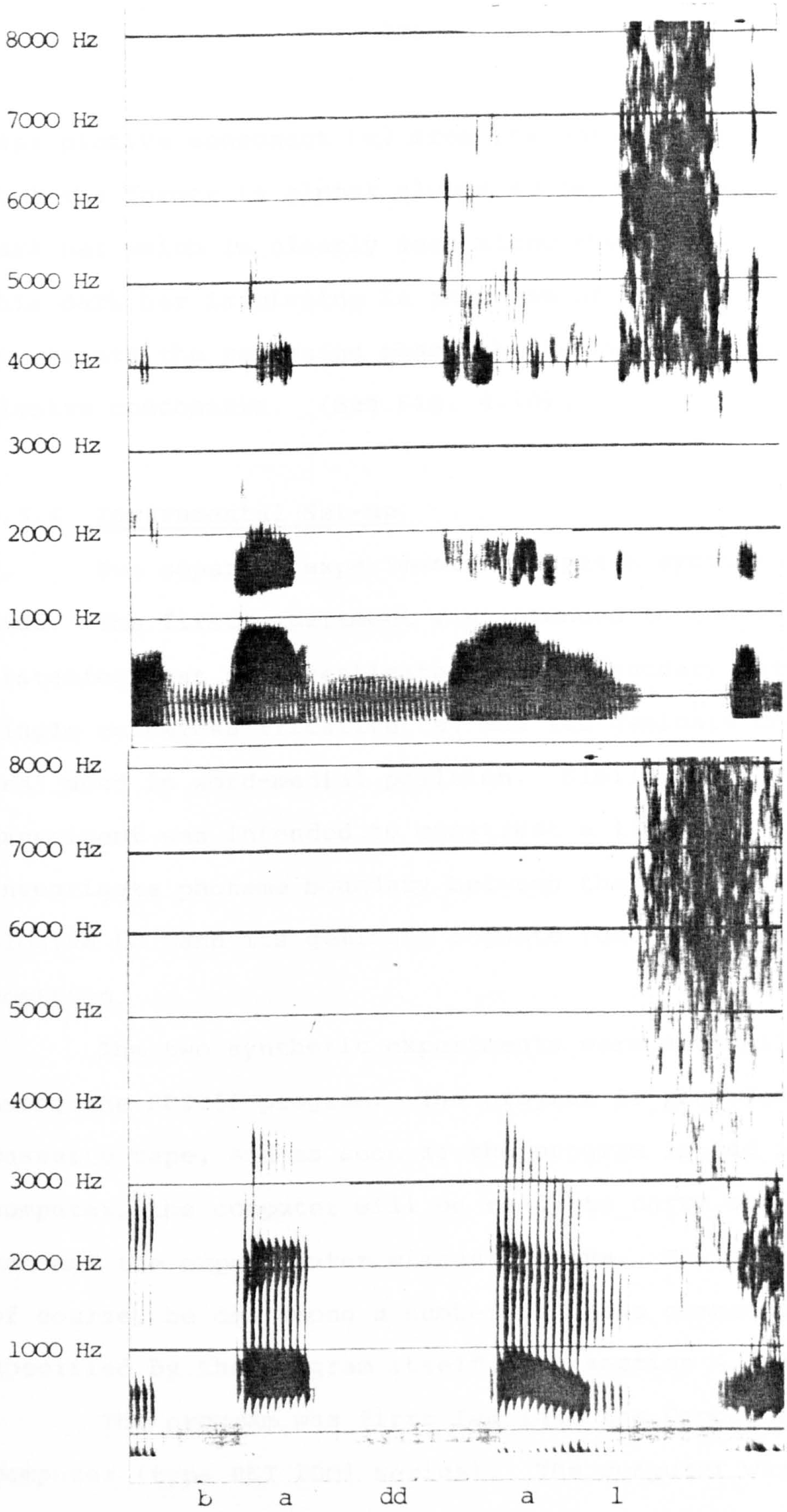


Fig. (4.10) Two visual displays (spectrograms) for the word /baddal/ in the carrier sentence /kitbi baddal sit marraat/. Above: as said by a real speaker; below: synthetic speech made by controlling the sound patterns in a terminal-analog speech synthesizer driven by the Department's PET computer.

less plosive consonant [t] from its voiced partner [d] is that the former is almost always accompanied by a vertical dark bar which is clearly seen along the frequency scale. This dark bar is missing in the case of [d]; it normally represents the releasing phase or the burst of voiceless plosive consonants. (See Fig. 4.10).

4.3.4 Instrumental Set-up

Two separate experiments on speech synthesis were made. The first experiment was intended to construct a listening test to investigate phoneme boundary between the single voiceless fricative [s] and its geminate counterpart [ss] used in word-medial position. Similarly, the second experiment was intended to construct a listening test to investigate phoneme boundary between the single voiced plosive [d] and its geminate cognate [dd] in the same word position.

The two synthetic experiments were accomplished by using the SPS256 program. The program is recorded on a cassette tape, and as soon as the program is fed into the computer, the computer will be ready to carry out whatever command the experimenter wishes to give. The command should, of course, be one among a number of those commands which are specified by the program itself, see section 4.3.2.

The program was first fed into the Commodore microcomputer (type PET 2001 series). The computer was linked to a computalker speech synthesizer (model CT-1) by means of a set of fine wires. Another set of fine wires linked the computalker speech synthesizer to a 24 K memory expansion unit (model Plessey Memories, type No. Petite 24). Both the PET computer and the computalker speech synthesizer were linked via separate leads to high quality tape-recorder (type Ferrograph-Logic 7). The output from the computalker speech synthesizer was automatically recorded on a tape mounted on the Ferrograph. The automatic recording took place as soon as we typed in the command R on the computer. The tape used for recording the listening tests items was of high quality (type Ampex 1200). The recording speed of the tape was 7½", i.e. 19 cm/sec. The line knob on the Ferrograph was set at 6 for the first listening test, and at 9 for the second listening test. The two sets were kept constant while constructing the two listening tests. The master record knob on the Ferrograph was set at 10, i.e. maximum degree, for both tests. The volume knob was set at 3, and the upper record track was used.

While recording the synthesized output from the computalker speech synthesizer onto the tape mounted on the Ferrograph, initial and final durations of silent intervals were inserted at the end of the preceding recorded speech and the following one. The duration for the initial silent interval was 1 second, and that for the final silent interval was 2 seconds; thus making a total period of silence duration of 3 seconds between each two successive recordings. The 3 second period of silence duration was found very helpful while conducting the listening tests. During the preparation of the final versions of the two listening tests a bleep noise was purposely inserted at the beginning and end of each test. The same bleep signal was inserted at the end of every five synthetic utterances. The technique of inserting the bleep noise was of ultimate importance to help the subjects while doing their listening tests as well as not to let them go astray while making their ticks on the answer sheets. The frequency of the bleep signal was at 3 kHz when it was recorded. Fig. 4.11 illustrates the instrumentation used in the two experiments on synthetic speech.

4.3.5 Selection of Stimuli

Just like our stimulus items used in our acoustic investigation of Chapter Three, two oppositions involving different contoid types were chosen for our present experiments on speech synthesis. The two single consonants [s] and [d] versus their geminate counterparts [ss] and [dd] were involved. Each consonant occurred in medial position within a bisyllabic word of which the first syllable bears the primary stress. Four words, namely /ħasan/ versus /ħassan/ and /badal/ versus /baddal/, were selected for the two experiments. Each word was used in one and the same carrier sentence, i.e. /'kitbi——'sit mar'raat/ 'Write—— six times'.

Wide-band spectrograms were made from real speech using the 700 Spectrograph. The spectrograms represented the target words, viz. /ħasan/, /ħassan/, /badal/ and /baddal/ used in the constructed carrier sentence. Measurements (in msec) were made for the acoustic contents of each spectrogram

Fig. (4.11) **Block diagram of the instrumental set-up used in the speech synthesis experiments**

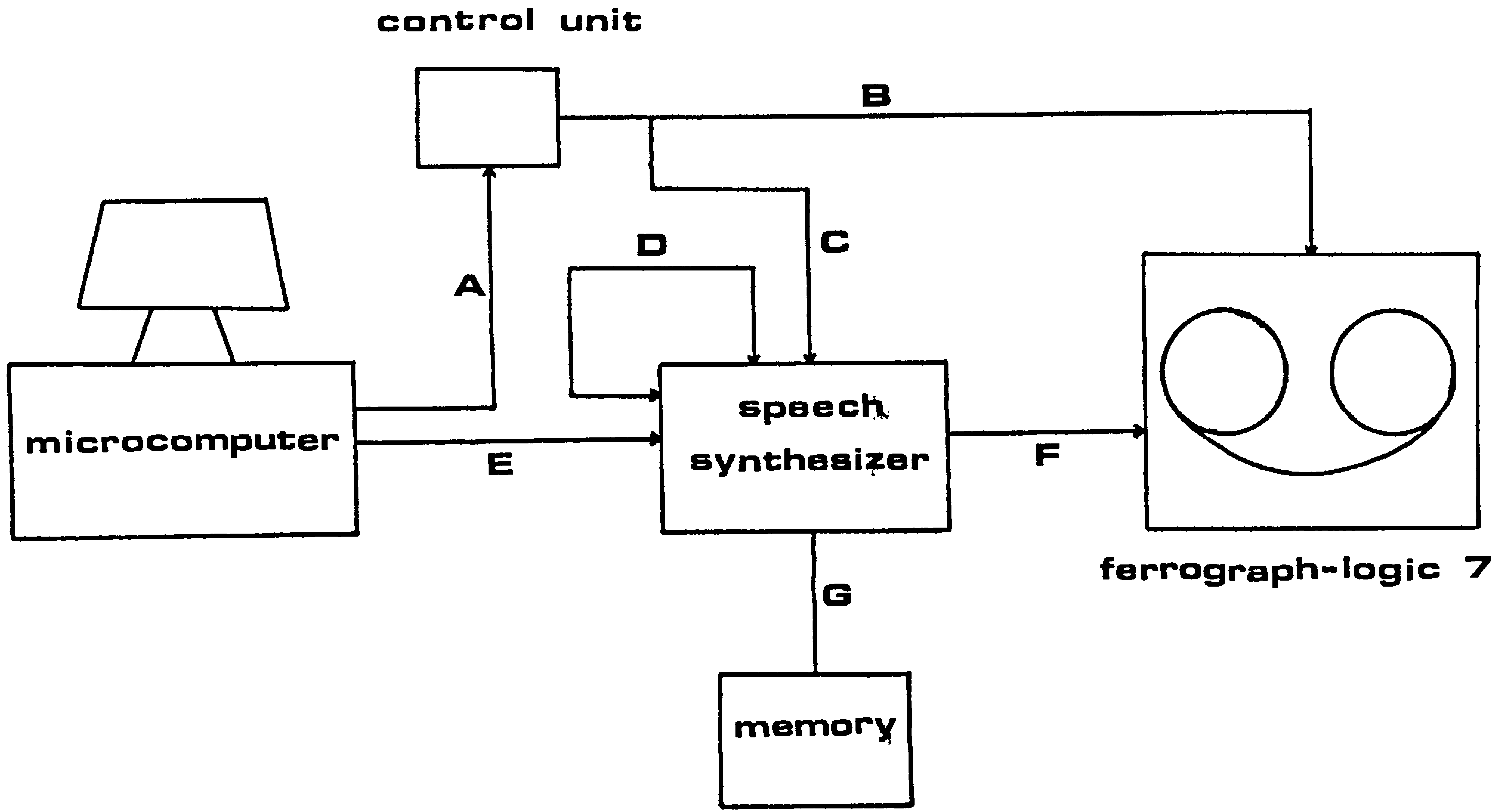


Fig. (4.11)

A	a lead linking the microcomputer (PET) to its control.
B	a lead linking the microcomputer (PET) to the Ferrograph-Logic 7 through PET control.
C	a lead linking the microcomputer (PET) to the Computalker Speech Synthesizer through PET control.
D	a set of fine wires going from and to the Computalker Speech Synthesizer.
E	a set of fine wires linking the microcomputer (PET) to the Computalker Speech Synthesizer.
F	a lead linking the Computalker Speech Synthesizer to the Ferrograph-Logic 7.
G	a set of fine wires linking the Computalker Speech Synthesizer to the memory expansion unit.

after deciding the number of the required points for every spectrographic segment. A total of 43 points was decided on the spectrogram representing the utterance /'kitbi 'ħasan 'sit mar'raat/; whereas a total of 41 points was decided on the spectrogram of the utterance /'kitbi 'badaḷ 'sit mar'raat/. In the case of [s] versus [ss], it was found that the duration of the shortest item [s] was 100 msec of which the PD value, representing the duration of the friction portion of [s], was only 45 msec; while the duration of the longest item [ss] was 280 msec of which the PD value was 225 msec. With regard to [d] versus [dd], on the other hand, it was found that the duration of the shortest item [d] was 75 msec of which the PD value, representing the duration of the closure portion of [d], was only 20 msec; whereas the duration of the longest item [dd] was 255 of which the PD value was 200 msec.

In both experiments ten gradual steps were made, beginning with the duration value of the shortest item and ending with the duration value of the longest item. Therefore, eight equal steps each of 20 msec increment were inserted between the two extremes, i.e. between the shortest duration value and the longest duration value.

In order to establish phoneme boundary between the single voiceless fricative [s] and its geminate cognate [ss], the shortest PD value was 45 msec and the longest was 225 msec. A list of the ten PD values was written out five times and, thus, making a total of 50 items. Each item, representing a PD value, was then written on a separate card. The fifty cards were carefully shuffled so as to arrange the fifty PD values in random order. To these fifty cards another five cards were added. The extra five cards contained five dummy items covering minimum, maximum and three intermediate values. Consequently, they raised the total number to 55 items.

The same procedure was also followed in establishing phoneme boundary between the single voiced plosive [d] and its geminate counterpart [dd], yet with one major exception. Here, the shortest PD value was 20 msec and the longest PD value was 200 msec. Similarly, eight equal steps each of 20 msec increment were added between the PD extreme values, i.e. the shortest value and the longest value.

4.3.6 Subjects

Ten male unpaid volunteers acted as subjects in each listening test. They were postgraduate students at Leeds University. All subjects were native I.C. Arabic speakers. They either came from the capital Baghdad, which lies in the central part of Iraq, or they came from the city of Basrah which lies far in the southern region of the country. Very few of the subjects came from places lying around the two big cities. They were all brought up educationally in either of the two cities, and they got their first higher degrees in one of their universities. They were all right-handed and reported no history of speech or hearing impairments. They were aged between 26-35 years.

Most of the subjects participated in both listening tests.

The other few apologized for not being able to sit for the second listening test either because they were entirely busy with their research projects or because they found the time to sit for the second test inconvenient. Therefore, they were replaced by other volunteers.

In as far as the phonetic experience is concerned, only one of the subjects, namely (AM), was a postgraduate student at the Department of Linguistics and Phonetics and he had had some experience in phonetics. The rest were, phonetically speaking, completely naive. None of the subjects had ever heard any synthetic speech before taking part in the present experiments.

4.3.7 Listening Test Conditions

The two listening tests were conducted in two separate sessions. The time between the first session and the second one was five weeks. Both listening tests were carried out in the Department Studio where the subjects were entirely cut off from any external noise or disturbance. The recorded items for the listening tests were played back on the Ferrograph-Logic 7 tape-recorder, and high quality headphones were used (type Super K, S.G. Brown Ltd.). The output volume level on the Ferrograph was set at 5 for the frame sentence, and at 3 for the bleep noise. This volume level was kept unchanged for all subjects while carrying out their listening tests.

Before starting the tests, the subjects were requested to meet together at one of the Department classrooms where they were instructed on how to do the tests. We reckoned that some of them would find it difficult to understand the instructions if they were reported in English; therefore, the Arabic language was used so as to make matters clearer and easier. Interpreting the instructions into English, they were as follows:

> You are going to hear a series of synthesized utterances spoken in everyday Iraqi Arabic. Each utterance consists of one and the same frame sentence which will either be heard as 'kitbi ħasan sit marraat' or 'kitbi ħassan sit marraat'. If you hear the former sentence, please mark the column on your answer sheet under the word 'ħasan' with a tick. If you hear the latter, mark the column under 'ħassan'. If you cannot decide between 'ħasan' and 'ħassan', you may guess and a tick *must* be placed under one of them.
>
> There are 55 items on the recorded test. The test will begin and end with a bleep noise. The same bleep signal will be heard at the end of every five utterances. This will remind you of the termination of the preceding five utterances as well as the initiation of the other following five.

The same instructions were given for the second listening test with the exception that the two words 'badal' and 'baddal' replaced the two words 'ħasan' and 'ħassan', respectively.

The subjects were requested to sit for the two tests in turn inside the studio. There, each subject was supplied with the appropriate answer sheets each time. Some subjects demanded a preliminary listening for the first five items; others needed extra copies of the answer sheets because they thought that they had made certain mistakes while identifying the test words. The unwanted copies of their answer sheets were cancelled and destroyed later. On the back side of the answer sheets, the subjects were informed to write down their full names, date and place of birth, and their field of specialization. In general, it took each subject about 10 minutes to do each listening test.

4.4 Statistical Test Used

The results of the listening tests in the present experiments on speech synthesis were subjected to a statistical treatment to determine, in terms of probability, whether the observed perceptual differences of cross-over points between the two independent samples, namely the fricatives [s] versus [ss] and the plosives [d] versus [dd], signify that the populations sampled are themselves really different or whether these differences exist merely because of chance factors. Consequently, this will enable us to determine whether "the observed difference is within the range which could easily occur by chance or whether it is so large that it signifies that the two samples are probably from two different populations." (Siegel, 1956, p.2).

In social sciences, such as those of linguistics and phonetics, the first step in the decision-making procedure is that the researcher has to state the null hypothesis (Ho for short) which is a hypothesis of no difference. The ultimate aim of formulating Ho is to be rejected, and if it is rejected, the alternative hypothesis (Ha for short) may, then, be accepted. Siegel (op.cit.) states that Ha is the operational statement of the experimenter's research hypothesis. The research hypothesis is the prediction derived from the theory under test. If we presume that the research hypothesis in our present experiments predicts perceptual differences in cross-over points between [s] versus [ss] on the one hand, and [d] versus [dd] on the other, then our null hypothesis must be that no such difference exists.

In determining which statistical test is most appropriate to choose for analyzing a particular set of research data, the researcher must keep in mind that the power of a statistical analysis is partly a function of the statistical test employed in the analysis. In addition to power, Siegel (op.cit.) adds other considerations which enter into the determination of which statistical test the researcher must choose. They may also include "the manner in which the sample of scores was drawn, the nature of the population from which the sample was drawn, and the kind of measurement or scaling which was employed in the operational definitions of the variables involved, i.e. in the scores." [p.18].

For the purpose of evaluating the reliability and the significance level of the results obtained in the experiments of this chapter, the Mann-Whitney U-test, which is regarded as one among the most powerful of the non-parametric tests, has been chosen for either rejecting or accepting Ho. Our choice for this statistical test has been based on the fact that the probability statements obtained from most non-parametric statistical tests are exact probabilities regardless of the shape of the population distribution from which the random sample was drawn; besides non-parametric tests simply take account of the rank order of the scores and allow us to test the more general hypothesis that one set of scores tends to be higher, or lower, than another set.[1] Miller (1975) remarks that by finding the sum of the ranks of one of the samples "the Mann-Whitney test allows us to determine the probability that a given separation between the ranks of the two samples could have arisen by chance." [p.82].

Since in a 1-tailed test the region of rejection is located entirely at one end of the sampling distribution, and in a 2-tailed test the region of rejection is located at both ends of the sampling distribution, a 2-tailed test has been used in this study. This is because we cannot be so specific and predict whether differences between fricatives and plosives, as far as the values of their cross-over points are concerned, are significant, and it is unwise to mix the two tests when the difference is expected to go both regions for other phonetic categories. Furthermore, with a 2-tailed prediction, i.e. non-directional hypothesis, we will analyze the data whichever direction the difference happens to be in.

---

1. Differences between parametric and non-parametric statistical tests will be discussed in section 6.11 of this study.

It is the researcher who must specify a level of significance at which Ho is rejected after choosing the appropriate statistical test for his study. He determines the significance level related to his findings by stating that Ho will be rejected at that level. The significance level is simply the probability of making a type 1 error, i.e. "when we decide that the independent variable had an effect on the dependent variable when it did not have." (Robson, 1973, pp.34-35). Miller (op.cit.) states that to specify a certain level of significance is essentially an arbitrary matter; a matter of convention rather than basic principle. Most experimental psychologists choose a significance level of 0.05. However, in certain circumstances we may even set ourselves a far more demanding level of significance so as to be that much more confident in any findings that reach our chosen level of significance. The lower the probability set for the significance level the greater the chance of making a type 2 error, i.e. "when we conclude that the independent variable had no effect on the dependent variable when in fact there was a genuine relationship." (Robson, op.cit., p.35).

In so far as the results of our present study are concerned, two levels of significance for rejecting Ho have been chosen. The first level is $p \leq 0.01$, where the differences between single and geminate consonants will be referred to as statistically highly significant. The second level is $p > 0.01$ and $p \leq 0.05$, where the differences will be referred to as statistically significant. Therefore, differences where $p > 0.05$ will be referred to as statistically non-significant. In the text, the first and second levels of significance will be marked by the signs (**) and (*), respectively. The non-significant level will be left unmarked.

4.5 Analysis of Results

4.5.1 Method Used for Examining Results

The items in the tests were set in random order, and for analysis of the results it was necessary to put the subjects' responses back into their original order. It was decided to use the computer to do this more efficiently. The computer was given the sequence of random numbers that had been used for randomizing the items, and then all the responses of the subjects were typed in.

For each item the card number minus 5 was typed. That is, the first 5 dummies were excluded and out of the 55 items that were recorded on the subjects' answer sheets the data for only 50 items were fed into the computer for each subject. Number '1' was typed in front of the items where the subjects indicated on the answer sheets that they had heard /ħasan/ in the first listening test, and /badal/ in the second listening test. At the same time, number '2' was typed opposite to the items where the subjects indicated that they had heard /ħassan/ and /baddal/ in the first and the second listening tests, respectively.

For the purpose of displaying the results, it was decided to present the subjects' responses in the form of a graph showing, for each duration, what percentage of a subject's responses was for single and what percentage was for geminate. In all cases the identifications start with complete separation (100% and 0%), and the two lines representing single and geminate scores cross each other approximately in the centre of the graph. It was felt that the duration value at which the traces crossed was important; this is usually called the 'cross-over point'.

4.5.2 Identification of Phoneme Boundary

It has been stated in the previous section that for both listening tests the individual and the pooled responses of all subjects were made visible in the form of graphs representing the subjects' identification of the synthetic speech stimuli as [s] or [ss], and as [d] or [dd]. On the graphs, the red line indicates the identification of either [s] or [d]; while the green line indicates the identification of either [ss] or [dd]. The graphs also show the percent judgments of the identification of the above synthetic speech items plotted against duration (in msec).

Since six of the subjects participated in both listening tests and the others in either of the two; therefore, our analysis of the results will first begin with those who took part in both tests and then we shall individually tackle the results of those listeners who sat only for either the first or the second listening test.

4.5.2.1 Results of Subjects who Participated in Both Listening Tests

Figs. 4.12 and 4.13 show how subject HJ assigned the boundary between the single fricative [s] and its geminate cognate [ss], and between the single plosive [d] and its geminate counterpart [dd]. It is immediately apparent from the two figures that the boundary between the two sounds of both classes is reasonably sharp for this subject. It lies at about 190 msec of consonant duration. Before 160 msec of duration the two singles [s] and [d] were fully identified; whereas the two geminates [ss] and [dd] were distinguished after 220 msec.

Similarly, Figs. 4.14 and 4.15 indicate how subject AR identified the boundary between the same two fricatives [s] and [ss], and the same two plosives [d] and [dd]. For this subject the two singles [s] and [d] were clearly distinguished before 180 msec of consonant duration and their geminate partners [ss] and [dd] were identified after 220 msec. However, for him the cross-over point representing the boundary between [s] and [ss] lies at about 197 msec of consonant duration, and that between [d] and [dd] lies at about 192 msec. It is also quite obvious from these figures that the phoneme boundary between the sounds of the two classes is reasonably sharp for this subject.

For subject SA the boundary between [s] and [ss] lies at about 189 msec of consonant duration. For him the single fricative [s] was identified before 140 msec and the geminate [ss] after 200 msec of duration. If compared with

Figs. 4.12 and 4.13, Fig. 4.16 is clearly less sharp in so far as the phoneme boundary is concerned. Between 160 msec and 180 msec of consonant duration the subject showed uncertainty of phoneme identification; 80% of judgment was for the single [s] and 20% was for the geminate [ss]. On the other hand, a cross-over point indicating the boundary between the two plosives [d] and [dd] lies at about 176 msec of consonant duration. [d] was identified before 160 msec, and [dd] was identified after 220 msec. A region of uncertainty was, to a certain extent, shown at 200 msec of consonant duration, see Fig. 4.17.

In Fig. 4.18 subject AA assigned the boundary between the single [s] and its geminate [ss] at about 183 msec of duration. It is quite evident from this figure that the identification of the phoneme boundary between the two consonants is reasonably sharp. [s] was identified before 160 msec; while [ss] was distinguished after 200 msec of consonant duration. As for the boundary between the single plosive [d] and its geminate partner [dd], a cross-over point between the two consonants lies at about 208 msec of duration. Fig. 4.19 shows that the identification boundary between [d] and [dd] is less sharp than that between [s] and [ss] for the same subject. The non-geminate [d] was fully identified before 160 msec; whereas the geminate [dd] was clearly recognized after 220 msec of consonant duration. In contrast with Fig. 4.16 of subject SA, this subject also showed uncertainty of phoneme boundary identification between 180 msec and 200 msec of consonant duration; 80% of judgment was for the single [d] and 20% was for the geminate [dd].

Subject IM assigned a cross-over point indicating the identification of phoneme boundary between the single [s] and the geminate [ss] at about 176 msec of consonant duration. [s] was well identified before 140 msec, and [ss] was clearly distinguished after 200 msec. For the same subject the boundary between [d] and [dd] lies at about 172 msec of duration. He identified [d] before 160 msec, and [dd] after 200 msec. As illustrated in Figs. 4.20 and 4.21, the identifications of the phoneme boundary between the two consonants of either class are reasonably sharp.

Figs. 4.22 and 4.23 illustrate how subject AM assigned the boundary between the two fricatives [s] and [ss] on the one side, and between the two plosives [d] and [dd] on the other. [s] was well distinguished before 160 msec; whereas [ss] was clearly identified after 220 msec of consonant duration. Similar to that of Fig. 4.17, a region of uncertainty of phoneme identification between the two consonants is shown at 200 msec of duration, see Fig. 4.22. A cross-over point between the two fricatives [s] and [ss] lies at about 176 msec. On the other hand, this subject showed a reasonably sharper identification of phoneme boundary between the two plosives [d] and [dd]. The boundary lies at about 177 msec of consonant duration. He clearly identified the non-geminate [d] before 160 msec and the geminate [dd] after 200 msec.

4.5.2.2 Results of Subjects Who Participated in Either the First or the Second Listening Test

In Fig. 4.24 subject TK identified the boundary between the single fricative [s] and its geminate counterpart [ss] at a cross-over point lying at about 189 msec of consonant duration. He clearly distinguished the non-geminate [s] before 150 msec, and the geminate [ss] after 200 msec. Just like subject SA (Fig. 4.16), this subject showed uncertainty of phoneme identification in the region that lies between 160 msec and 180 msec of consonant duration. A judgment of about 80% is shown for identifying the single fricative [s], and the geminate cognate [ss] judged at about 20%. In the meantime, a less degree of uncertainty of phoneme identification was shown by subject AH at 160 msec of consonant duration, see Fig. 4.25. This subject assigned a boundary between [s] and [ss] at a cross-over point having the durational value of 184 msec. For him the single [s] was fully identified before 140 msec; while the geminate [ss] was clearly distinguished after 200 msec of duration.

Fig. 4.26, on the other hand, illustrates how subject FA assigned the boundary between the same two consonants, i.e. [s] and [ss]. An inspection of this figure will reveal the fact that the boundary between the two consonants is apparently sharp. He identified the single [s] before 140 msec, and the geminate [ss] after 180 msec of consonant duration. The boundary between them lies at about 163 msec. A similar sharp configuration representing the boundary of phoneme identification was shown by subject HK. However, an examination of Fig. 4.27 will show that the results produced by this subject is in remarkable agreement with that of subject AR (Fig. 4.14) in the way they assigned the boundary between the single [s] and its geminate counterpart [ss]. For both subjects [s] was clearly identified before 180 msec, and [ss] was fully distinguished after 220 msec of consonant duration. What differentiates between the two subjects is that the boundary between the two consonants lies at about 192 msec for subject HK, and it lies at about 197 msec for subject AR.

In regard to the identification of the single plosive [d] and its geminate cognate [dd], Fig. 4.28 shows how subject RK assigned a boundary between the two sounds at a cross-over point lying at about 184 msec of consonant duration. For this subject [d] was apparently distinguished before 160 msec, and [dd] was well identified after 220 msec. It is quite evident from this figure that the boundary between the two sounds is reasonably sharp. For subject MA the boundary between the same two consonants occurred at a cross-over point that lies at about 183 msec of duration. He distinguished the single [d] before 160 msec, and the geminate [dd] after 200 msec. It is also obvious from Fig. 4.29 that the boundary between the two sounds, for this subject, is reasonably sharp. Similarly, Fig. 4.30 shows how subject FN clearly identified [d] before 160 msec, and [dd] after 200 msec of consonant duration. He assigned a boundary between the two sounds at 172 msec. Fig. 4.30 illustrates that the boundary between [d] and [dd] is also reasonably sharp for this subject.

In contrast with the results obtained by the other subjects, a rather extreme value of consonant duration indicating the boundary between the single [d] and its geminate counterpart [dd] was shown by the two subjects AA and CG. Fig. 4.31 shows how the latter assigned the boundary between these two sounds at a cross-over point lying at about 203 msec of consonant duration. For him [d] was fully identified before 180 msec, and [dd] was well distinguished after 220 msec. The boundary between the two consonants has a reasonably sharp form.

The values of cross-over points resulting from the subjects' responses in the two listening tests are presented in tables 4.1 and 4.2. Table 4.1 shows the durations of cross-over points (in msec) of the first listening test; whereas table 4.2 illustrates the durations of cross-over points (in msec) of the second listening test.

When the overall values between the two classes of consonants were subjected to the evaluation of Mann-Whitney U-test to find out the significance level, the results indicated that the differences between the values of the cross-over points of [s] in /ħasan/ versus [ss] in /ħassan/ and of [d] in /badal/ versus [dd] in /baddal/ are statistically non-significant, with p value at 0.9699. Table 4.3 shows the differences of values of cross-over points (in msec) between the fricatives [s] and [ss] on the one side, and between the plosives [d] and [dd] on the other.

Table (4.1) Durations of cross-over points (in msec) of test (1) 'ħasan and ħassan'.

Subjects	Cross-over points
HJ	190
AR	197
SA	189
AA	183
IM	176
AM	176
TK	189
AH	184
FA	163
HK	192
Mean	184
S.D.	9.9

Table (4.2) Durations of cross-over points (in msec) of test (2) 'badal and baddal'.

Subjects	Cross-over points
HJ	190
AR	192
SA	176
AA	208
IM	172
AM	177
RK	184
MA	183
FN	172
CG	203
Mean	186
S.D.	12.5

Table (4.3) Differences of values of cross-over points (in msec) between fricatives and plosives.

Tokens	[s] ~ [ss]	[d] ~ [dd]
1	190	190
2	197	192
3	189	176
4	183	208
5	176	172
6	176	177
7	189	184
8	184	183
9	163	172
10	192	203
Mean	184	186
S.D.	9.9	12.5
U	49	
p	0.9699	

4.6 Discussion

In the previous chapter of our present study, evidence has been presented to indicate that geminate consonants are considerably longer than non-geminate consonants, whether they occur word-initially or word-medially. A geminate fricative or plosive was shown to be twice or three times longer, in terms of duration, than its single counterpart when occurring in the same phonetic context. These acoustic findings have been later proved to be perceptually true.

Our experiments on speech synthesis have confirmed the findings referred to in our acoustic experiment. The two listening tests conducted in our two experiments of this chapter have indicated that listeners could easily distinguish geminate consonants from their non-geminate partners in synthetic speech when the duration of the consonant was varied. A fact that has been apparently illustrated by means of a number of graphs representing the subjects' responses in the two listening tests. The graphs have also indicated that some subjects showed uncertainty in identifying a single consonant from its geminate cognate. This uncertainty of phoneme identification was generally between the regions 160-200 msec.

With reference to the above statement, one would like to know whether the cross-over point was affected by our choice of the shortest and longest duration values. More specifically, whether the cross-over point would move 20% higher if we started with a minimum or a maximum duration of 20% longer than the one we used in our present experiments. If this was really the case, then one would also like to know the extent to which the listener's cross-over point would be shifted if we kept the s~ss continuum durations and the d~dd continuum durations constant but made the carrier sentence's segment durations shorter or longer. To answer these questions we would, undoubtedly, need to keep on doing more and more experiments around this area such as, for instance, taking a duration continuum of some non-speech sound, in the same way as that suggested by Liberman et al. (1957, p.367), and see how sharply the listeners' long versus short judgments cross over. Such an experiment, very clearly, lies beyond the scope of our present investigation.

However, it is our aim now to show whether the listeners' identification of single consonants from their geminate partners occurs categorically, i.e. they are always and only perceived as tokens of a particular phonetic type. (Studdert-Kennedy, 1973). This makes it necessary, at this stage, to refer to those studies[1] that dealt with the phenomonon of 'categorical perception' and discuss whether the data we have obtained from our two experiments on speech synthesis are an example of it.

4.6.1 Categorical Perception

In section 4.2.1 it has been explained that early works dealing with speech synthesis showed that an approp-

---

[1] Some of these works have already been discussed in section 4.2.1.

riate procedure for defining the acoustic properties of a phoneme segment was to construct tokens of opponent categories, distinguished on a single phonological feature, by varying a single acoustic parameter along a continuum. For example, changing /ba/ to /da/, /da/ to /ta/, /ba/ to /ga/, and so on. (Studdert-Kennedy, 1979). When these tokens were presented to listeners for identification, they tended to identify any particular stimulus in the same way every time they heard it. It was also found that when the listeners were instructed to discriminate between adjacent tokens, they tended to respond very badly if they assigned the two tokens to the same category; yet they responded very well if they assigned them to different categories despite the fact that the acoustic distance between the tokens was identical in the two categories.

Categorical perception assumes that listeners can discriminate two sounds only to the extent that they can identify them as different. Since discrimination is thus based exclusively on the phonetic category labels, the obtained discrimination function must be non-significantly different from that predicted from the identification data. Ferrero et al. (1982) remarked that when the discrimination function may be predicted from the identification function, that is, when the discrimination is based exclusively on the phonetic category labels, as in the case for plosive consonants, the perception is said to be 'categorical'. On the other hand, when the obtained discrimination function is better than the predicted, it means that the subjects discriminate differences which they cannot identify and perception is said to be 'continuous'.

In an earlier paper, Studdert-Kennedy et al. (1970) stated that categorical perception "refers to a mode by which stimuli are responded to, and can only be responded to, in absolute terms." [p.234]. They reported that successive stimuli which were drawn from a physical continuum were perceived as members of discrete categories and not as forming a continuum. This is to say, they were identified independently of the context in which they existed. Carney et al. (1977) found that a subject's discrimination performance was determined by phonetic judgments rather than by auditory judgments. They commented that "since identification functions can be made arbitrarily steep by appropriate scaling of the stimulus dimensions, the second condition, that pertaining to discrimination performance, is the more restrictive." [p.961]. They also added that the condition implied a true discrimination threshold; within-category differences were subthreshold and were never perceived, whereas between-category differences were suprathreshold and were always perceived. This is, probably, why the existence of such thresholds is not universally accepted; while the threshold like properties of categorical perception perhaps underlie the controversy surrounding this concept.

For the Haskins Laboratories researchers the term categorical perception is primarily restricted to the perception of speech stimuli and that it is typical with consonants rather than with vowels. Concerning the voicing feature, their studies on speech synthesis indicate that the distinction between voiced and voiceless sounds is perceived in a categorical mode. The listeners tend to respond to the speech sounds that will be identified as a single phonetic segment. Liberman et al. (1961), for instance, found that synthetic speech stimuli which varied in steps that were acoustically equal through the range sufficient to produce the initial plosive consonants [d] and [t] were heard as members of discrete categories. When listeners were exposed to hear the neighbouring pairs of these stimuli in an ABX procedure[1] for a discrimination test, they were able to discriminate stimuli extracted from different phonetic categories, but unable to discriminate between those stimuli extracted from within the same phonetic category albeit the acoustic differences were comparable in both cases.

The distinction between categorical and continuous, i.e. non-categorical, modes of perception has played a significant role in theoretical discussions of speech perception. Pisoni and Lazarus (1974) reported that some researchers, e.g. Liberman (1970); Liberman et al. (1967); Stevens and House (1972), have been led to propose that the perception of speech may involve specialized perceptual processes that are basically different from the process involved in the perception of auditory stimuli unlike speech. They contended that when listenrs were exposed to certain classes

---

[1] Also see section 4.2.1 for further details about the ABX procedure for a discrimination test.

of speech sounds their ability to identify and discriminate between them on an auditory basis was somehow constrained. Listeners did not perceive these stimuli as isolated acoustic events but rather responded to them with reference to their linguistic knowledge.[1]

Liberman (op.cit.) stated that the phenomenon of categorical perception is an unusual result to get in psychological experiments with nonspeech sounds, and stimuli that vary along continua of physical nature, such as intensity and/or frequency, are normally perceived in a continuous mode. Pisoni and Lazarus (op.cit.) commented on this point saying that "the discrimination functions are usually monotonic with the physical scale, and the listeners can discriminate many more stimuli than they can reliably identify in absolute terms." [p.328].

For Pastore et al. (1977) the term categorical perception refers to the apparent responding to stimuli only in absolute terms. They suggest that categorical perception involves the following four basic factors:

a) distinct labelling categories separated by a relative sharp boundary;

b) regions or 'troughs' of chance performance for discriminating stimuli within the same category;

c) a peak in the discrimination function at the category boundary; and

---

[1] Cf. Studdert-Kennedy, 1974, p.2360.

d) a close correspondence between the actual discrimination performance and performance predicted on the basis of the assumption of absolute categorization along the given continuum.[1]

Liberman et al. (op.cit.), depending on a number of experimental studies involving the perception of speech and non-speech stimuli, proposed a model for speech perception which assumes that phonological units are identified by a special decoder which refers the incoming speech sounds to the invariant neuromotor commands controlling the articulatory muscles. This model and others[2] have been dubbed 'the motor theory of speech perception'.

This theory was comprehensively criticized by Lane (1965) who questioned the appropriateness of the nonspeech control stimuli applied in a number of investigations dealing with the phenomenon of categorical perception. According to Lane the nonspeech control stimuli should be equivalent to the speech stimuli, and "in order to assess the degree to which the speech-discrimination functions reflect acquired similarity, acquired distinctiveness, or both, it is necessary

---

[1] Following Studdert-Kennedy et al. (op.cit.), Pastore (1976) suggested three major criteria which are specified for the phenomenon of categorical perception. They are: "First, the labelling function must be clear-cut, with relatively sharp boundaries between identified groupings. Second, distinct peaks in the discrimination function must appear at these boundaries. Finally, there must be good agreement between the obtained discrimination function and the discrimination function predicted on the basis of assumed absolute categorization. If the first criterion is met, this last criterion should include the presence of definite 'troughs' of chance performance in the discrimination function within categories." [p.253].

[2] For example, see Stevens and House (op.cit.), and other works by the researchers of the Haskins Laboratories.

to have a base line of discrimination which presumably reflects the discriminability of the speech stimuli 'in the row' - that is, before language learning has taken place." [p.283]. He pointed out that the control stimuli should be comparable to the speech stimuli, varying in the same acoustic properties and with the same constant parameters, so that the difference in discriminability of the speech and control stimuli might not be attributed to a difference in the psychophysical task. Lane also cited the fact that the discrimination functions obtained for speech stimuli are somewhat better than predicted on the basis of the labelling functions. He stated that the obtained discrimination functions lie, at all points, somewhat above those predicted, suggesting that the listener can discriminate between two speech stimuli slightly more often than he can identify them absolutely. Pastore (op.cit.) remarked that Lane drew our attention to the existence of small peaks in the discrimination functions for speech stimuli which did not match with categorical boundaries and therefore must not be 'a product of absolute categorization'.

Lane (op.cit.) rejected the claims of the Haskins Laboratories researchers for the motor theory of speech perception. He considered some methodological limitations of previous studies that supported the motor theory of speech perception and presented evidence indicating that the reported relations between identification and discrimination functions are not all unique to the perception of consonants but describe as well the perception of vowels and of entirely nonlinguistic stimuli. He believed that these findings with nonlinguistic stimuli militate against the postulation of a special perceptual mechanism for speech. As for the opinion put forward by the researchers of the Haskins Laboratories that perception of the consonants is more categorical than is perception of the vowels, Lane stated that "the same findings for the vowels stimuli also contravene the motor theory for this reason: the theory maintains that the perception of most consonants is discontinuous or categorical because they covertly evoke discontinous articulations whose discretely different sensory feedback controls the overt identifications, whereas the perception of vowels is continuous because the mediating articulations are continuous and provide similar sensory returns." [p.292]. He, then, concluded that there was considerable evidence to support the conclusion that the general form of labelling functions for consonant continua did not differ from that for vowel and nonlinguistic continua[1]; besides there was also evidence that the detailed form of labelling gradients for consonant and nonlinguistic continua did not differ appreciably. For him the relation between the identification and discrimination of consonants is also found with vowels and nonspeech stimuli, and that "these relations are attributable to the general paradigm for discriminative training and testing and that the postulation of a special perceptual mechanism for the discrimination of speech stimuli is unwarranted." [p.307].

---

[1] Cf. Studdert-Kennedy et al., 1970.

Lane's critical review of the motor theory of speech perception was opposed by Studdert-Kennedy et al. (op.cit.) who strongly expressed their disagreement with Lane's interpretations. They remarked that they could not agree with the implication that the degree of categorical perception did not differ from those several types of continua "since a sharp identification function is not a sufficient condition of categorical perception." [p.237]. They also confirmed a previous remark that subjects might learn to identify stimuli consistently if the stimuli were spaced widely enough along the continuum. Studdert-Kennedy and his co-workers criticized Lane's attempts to establish categorical perception of nonspeech stimuli by identification training. They stated that "while training may, for some subjects, increase distinctiveness between categories and produce a degree of categorical perception such as that found with vowels, we have no evidence for an approach through training to decreased distinctiveness within categories and the degree of categorical perception found with consonants." [p.242].

The basic disagreement between Lane and Studdert-Kennedy et al. (op.cit.) was that in the interpretation of the concept of categorical perception. While Lane believed that categorical perception was merely a consequence of discrimination training, Studdert-Kennedy and his colleagues were willing to entertain the hypothesis that it reflects some structurally determined process, adapted to the complex code that links the sounds of speech to the phonetic message they convey. They believed that the basic difference between plosive consonants and steady-state vowels was that "plosives are more complexly encoded in the sound stream and, therefore, more in need of a special speech-sound processor." [p.248].[1] They claimed that both experience and training were sufficient conditions for categorical perception or for other characteristics of perception in the speech mode, and that categorical perception would be one result of the operation of a special decoding device available to man as part of his species-specific capacity for language. They also pointed out that Lane did not deny that perception might be categorical, and added that he saw the phenomenon as quite general.

Recently a number of studies have demonstrated categorical perception with musical stimuli. Burns and Ward (1974), for example, have conducted an investigation of the perception of musical intervals by experienced musicians using the usual procedures related to categorical perception experiments. Identification and discrimination functions have been successfully obtained from points spaced equally along a physical continuum; that is to say "equal distances in cents along the logarithmic frequency scale defining musical pitch." [p.456]. The results obtained have shown that the musicians exhibit categorical perception, some to a degree approaching that shown in the perception of stop consonants.

The study by Cutting and Rosner (1974), on the other hand, has clearly demonstrated categorical perception for

---

1. Cf. Liberman et al., 1967.

nonspeech stimuli. The rise time of three types of stimuli was varied in 10 - msec steps between 0 and 80 msec. The three types of stimuli were: sawtooth wave, speech syllable, and sinusoid. The results were clear and equivalent both with the sawtooth and speech stimuli; while in the case of the sine waves somewhat similar discrimination results were obtained, and at the longer rise times the labelling categories were less well separated. The investigators also noted that categorical perception was typically demonstrated with complex auditory signals, i.e. plosive consonants, noise-buzz sequences, and sawtooth waveforms, by varying the onset times of the given stimuli which all exhibited categorical boundaries at onset times of approximately 20 msec. Pastore et al. (1977) stated that "this similarity would seem to imply the existence of a common causality for the categorical perception of stimulus-onset manipulations." [p.695].

Other researchers (e.g. Eimas, 1975; Kuhl and Miller, 1975; Morse and Snowdon, 1975) have extended their work and carried out a number of experiments which have indicated that young infants and even some animals such as chinchillas and rhesus monkeys might exhibit categorical perception. These studies implied that the phenomenon of categorical perception might even be more general, and was probably more basic, than implied by previous investigations in this field of speech perception.

4.6.2 Categorical Perception in Relation to Our Present Data

In the light of the results obtainable in the present experiments on synthetic speech, it is still difficult for us to prove whether these results reflect a genuine case of the phenomenon of categorical perception despite the fact that one may get the impression that they do so simply by having a glance at the graphs showing the responses of the subjects who participated in the two listening tests. It is a fact, however, that a really detailed study of the perception of single versus geminate differences would need a thorough examination, and experimental follow-up, of various factors. Among these factors is the most important question of the possibility of a categorical perception effect which we have already reviewed in the preceding section. In order to pursue this question, at least in later research work, it is very clear that we would need to conduct a parallel experiment using the same duration continuum to test for peaks in the discrimination function, or we would need to do further synthetic experiments aiming at illustrating non-categorical perception, i.e. non-linguistically significant differences, such as vowel duration or, for instance, word-initial [γ] versus [γγ]. The latter example is linguistically non-significant because gemination does not exist in this case in Arabic. Nevertheless, if this had been predominantly a perceptual/synthetic study, and if time had permitted, then we would have done so.

It is quite obvious that the above statement will lead us to the fact that if we want to consider whether the differences in duration between single and geminate consonants studied in our present experiments are within the range that a human listener can detect, then it would be a plausible idea to refer to the difference limens (DL's for short)[1], or as it is sometimes called the just-noticeable differences (JND's for short). It is believed that phoneme identification tests seem to be an excellent way of measuring DL's despite their remoteness from normal speech perception (Schouten, 1980; Nooteboom, 1981).

Works by Stott (1935), Henry (1948), Small and Campbell (1962), Ruhm et al. (1966), Chen (1970) and Huggins (1972) have copiously dealt with the DL's for duration. A considerable amount of work has been done aiming at establishing the Weber ratio, namely $\Delta T/T$, for various reference durations. The Weber ratio is "the ratio $\Delta T/T$, that is change in duration over reference duration." (Lehiste, 1970, p.11). According to Weber's hypothesis the ratio between the stimulus increment and the reference stimulus is a constant, and these constant ratios apply to all sense modalities.

Stott's (ibid.) experiment aimed at making an accurate and thorough discrimination of the constant or time-order errors which occur in comparing short tonal durations, i.e. intervals of time filled with continuous tone. He used the method of constant stimuli, and a complete stimulus presentation consisted of a standard

---

[1] The DL has been defined as "the smallest detectable change in a stimulus." See Moore, B.C.J. (1977). Introduction to the Psychology of Hearing. Macmillan Press Ltd., p.127 and p.302. For other relevant definitions, see Stevens and Davis, 1938, p.84; Denes and Pinson, 1963, p.114.

duration and a variable duration separated by an interval of silence of 1.5 sec. His data were for a 1000 cycle tone at an unspecified level with an interstimulus intervals of 1.5 sec.

Henry's (op.cit.) investigation, on the other hand, was concerned with the nature of the difference limens for duration of very brief sounds ranging from about 30 msec to nearly ½ sec in length, with some explanation of the effects on this limen of variations in stimulus quality and intensity. He established the Weber ratio and the absolute DL to be 0.196 and 21.56 msec respectively, for a comparable reference duration of 100 msec. The results of his three experiments indicated that the Weber ratios were not consistent. He remarked that by comparing the Weber ratio at short and long durations, there was confirmation of the trend for a smaller ratio at the longer duration. In so far as the Weber ratios and the DL's were concerned, his results showed clear differences from those reported by Stott. He believed that those differences were probably caused by using different experimental techniques and different instrumentation in the two studies.

Small and Campbell (op.cit.) tried to determine the differential threshold for duration and investigate discrimination at very short durations. They used a wide range of acoustic stimuli. The test signals were tones of 250 and 5000 Hz, and noise was presented at a 50-phon level. DL's were assessed across a range of reference duration between 0.4 and 400 msec, and at reference durations where comparisons were possible, the limens obtained by Small and Campbell were in the same order of magnitude as those reported by Henry (op.cit.) and by Stott (op.cit.). It was found that Weber's law did not hold across the reference durations used; the shorter durations yielding the larger limens. They also found no constant error when reference durations of 40 msec or longer were used in conjunction with inter-signal intervals of 0.2 sec or higher.

The Weber ratio and absolute DL's established by Ruhm et al. (op.cit.), Chen (op.cit.) and Huggins (op.cit.) were much smaller than those found in earlier investigations. Ruhm et al., for instance, reported a Weber ratio of 0.0260 or an absolute DL of 2.6 msec for a reference duration of 100 msec. Lehiste (op.cit.) commented on the apparent discrepancies in the DL's established by earlier studies as against those expounded in more recent investigations. She stated that "the research technique employed in the more recent study, was probably more conducive to testing the limit of the auditory sensitivity of the subjects." [p.12]. She also pointed out that it was perhaps a reasonable assumption that the DL's established by Ruhm et al. represented the limit of perceptibility under optimal conditions; whereas it appeared likely that in a speech condition, the JND's established by Henry and Stott might apply. It is, therefore, immediately evident that the Weber ratios and DL's established in Ruhm's study are much smaller than those found in other previous studies. Moreover, Ruhm et al. (op.cit.), using frequencies of 250 and

1000 Hz, found that signal frequency had no effect on the magnitude of DL's for duration. They discovered that sensation level affected the perception of differences in duration; a smaller limen was obtained at 50 dB sensation level than at 10 dB sensation level.

Chen (op.cit.) combined both Stott and Henry's findings, and compared them with his own data. He considered the mean vowel duration in each of the six different languages he studied as a reference duration equivalent to the reference duration in Stott and Henry's data. He also considered that the mean difference of vowel duration as a function of the voicing of the following consonant as equivalent to the ΔT. Fujisaki et al. (1975) found that the DL near a phoneme boundary for a two-syllable word spoken in isolation or in a carrier sentence was about 10 msec for vowels, fricatives and nasals. On the other hand, Klatt (1976) reported that such small values approached basic psychophysical limits for duration discrimination. Studies resulting in these small DL's, as Klatt pointed out, usually provide subjects with the opportunity to perceive individual segments, words or sentences repeatedly, thus permitting a stable reference pattern to be built up. After considering results from experiments offering less favourable conditions for discrimination, and perhaps more comparable with a normal speech hearing situation, Klatt concluded that changes of less than about 25 msec were perceptually of considerably less importance than changes above this duration, and that only changes of about 20% or more in duration served as significant perceptual cues.

Now, after all that has been discussed about establishing the DL's for duration, we are inclined to accept the suggestion put forward by Lehiste (op.cit.) that the range of the durations of speech sounds - usually from 30 to about 300 msec - the JND's in duration are between 10 and 40 msec. This fact has been quite recently confirmed by an investigation on vowel duration in Iraqi Arabic. Hassan (1981), following Lehiste's suggestion, finds that in all his mean differences of vowel durations none is below that range, and adds that some of them are even higher. He draws the conclusion that differences of duration for all the sequences hovering around or above the DL's line are likely to be perceived by the listeners.

Figs. (4.12-4.33)

Graphs of Identification Functions for Perception Experiments: /hasan ~ hassan/ and /badal ~ baddal/.

Fig. (4.12) /ħasan ~ ħassan/ Subject (HJ)

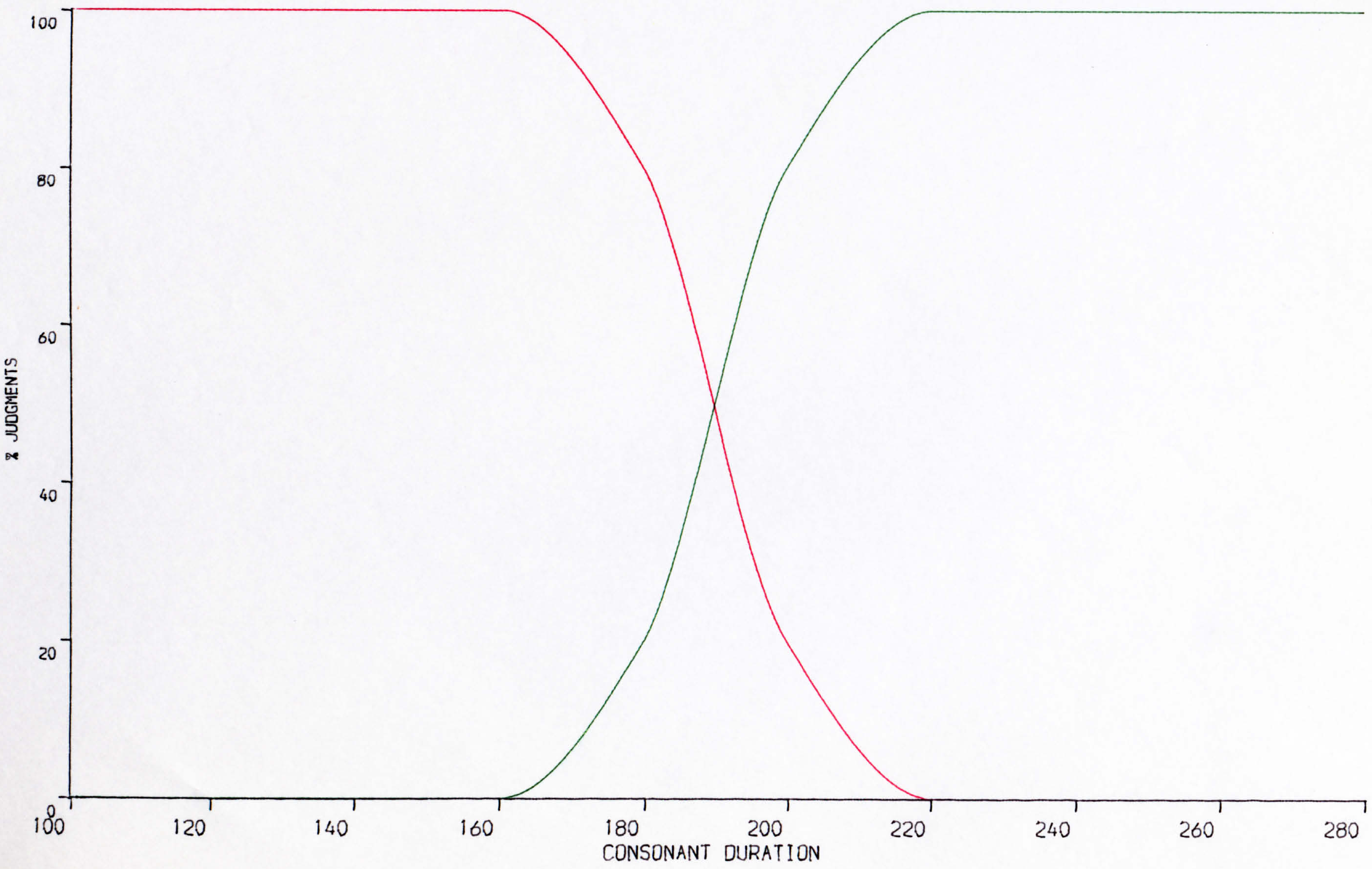


Fig. (4.13)

/badal ~ baddal/

Subject (HJ)

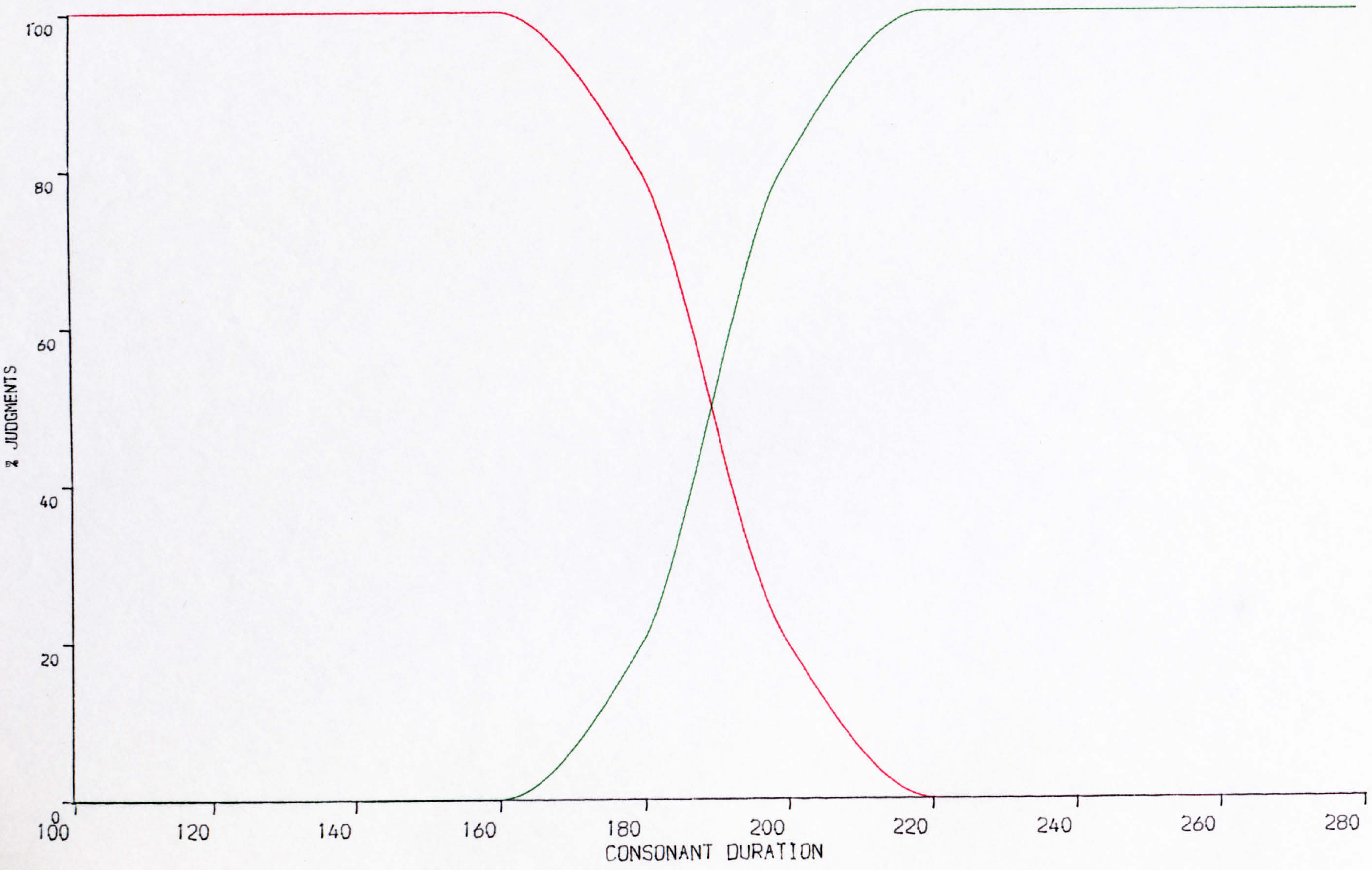


Fig. (4.14) /ħasan ~ ħassan/ Subject (AR)

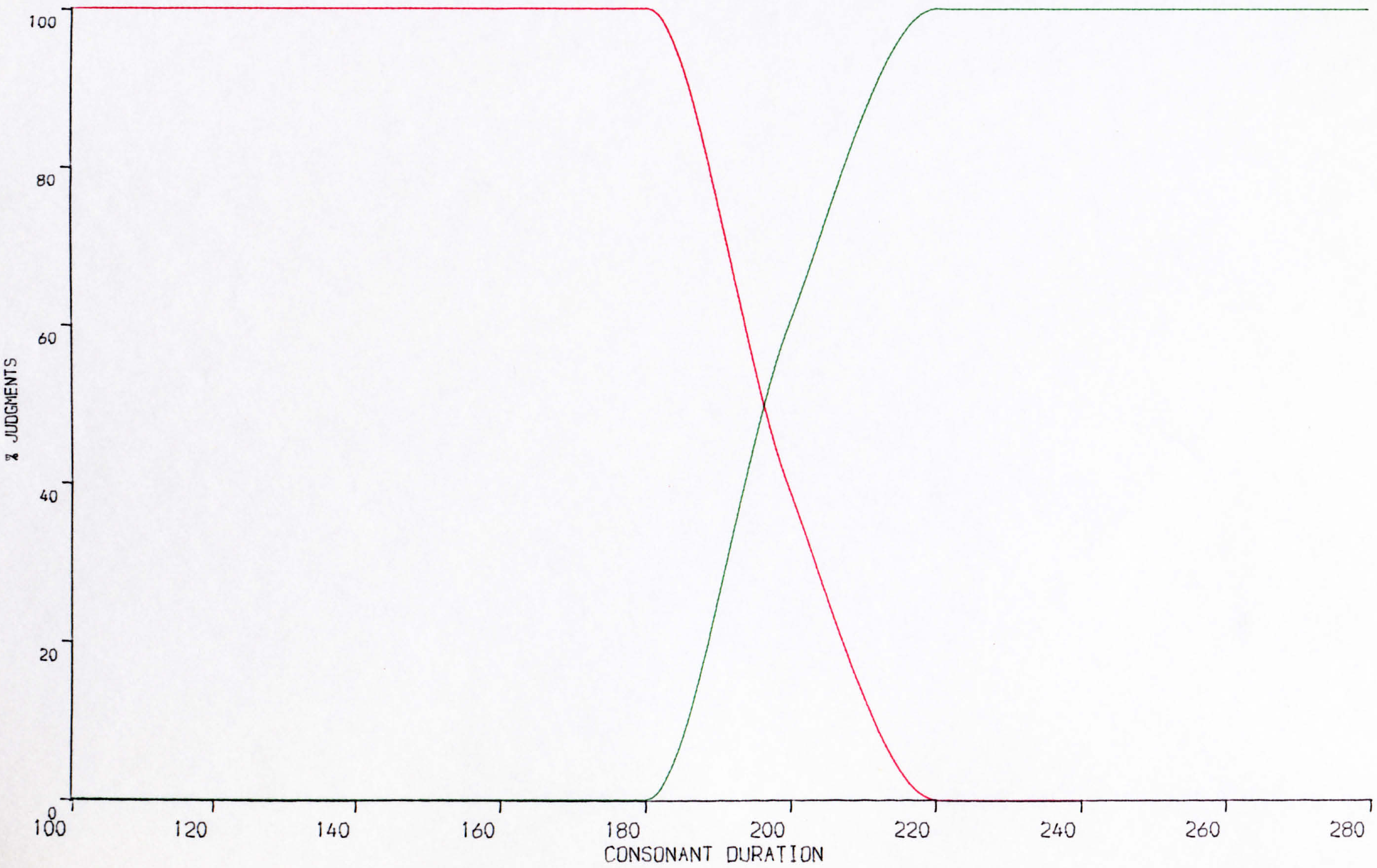


Fig. (4.15)

/badal ~ baddal/

Subject (AR)

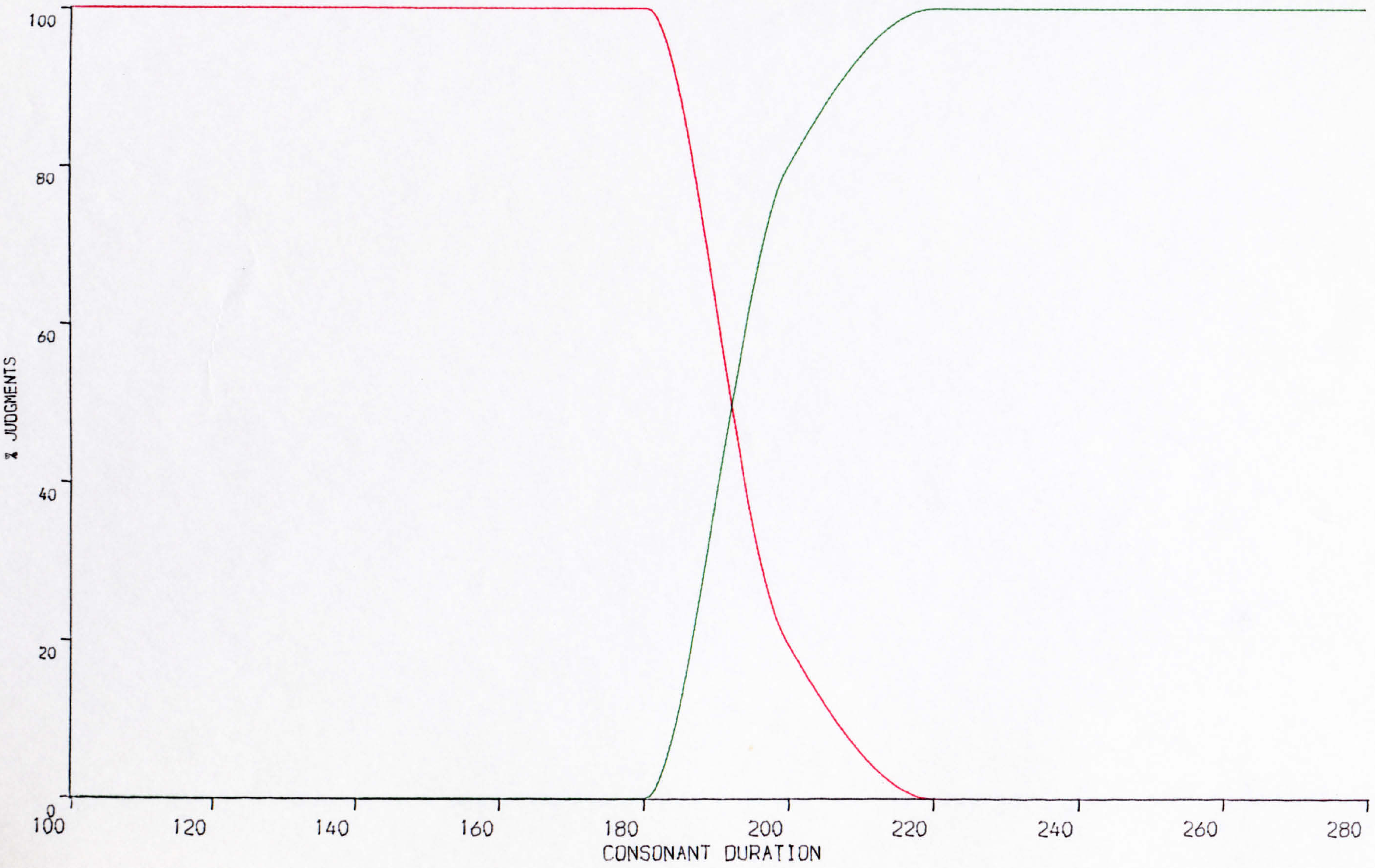


Fig. (4.16) /ħasan ~ ħassan/ Subject (SA)

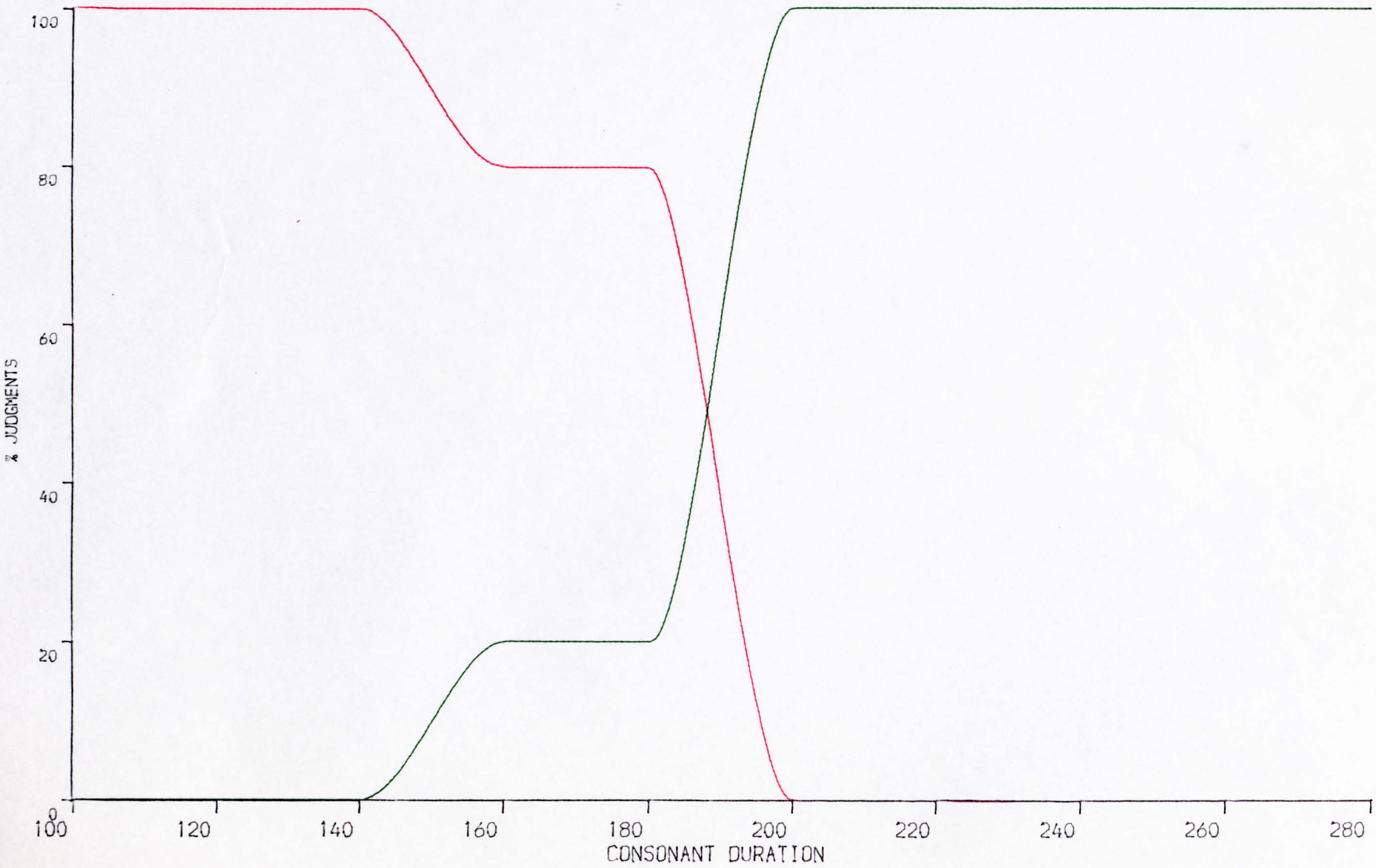


Fig. (4.17) /badal ~ baddal/ Subject (SA)

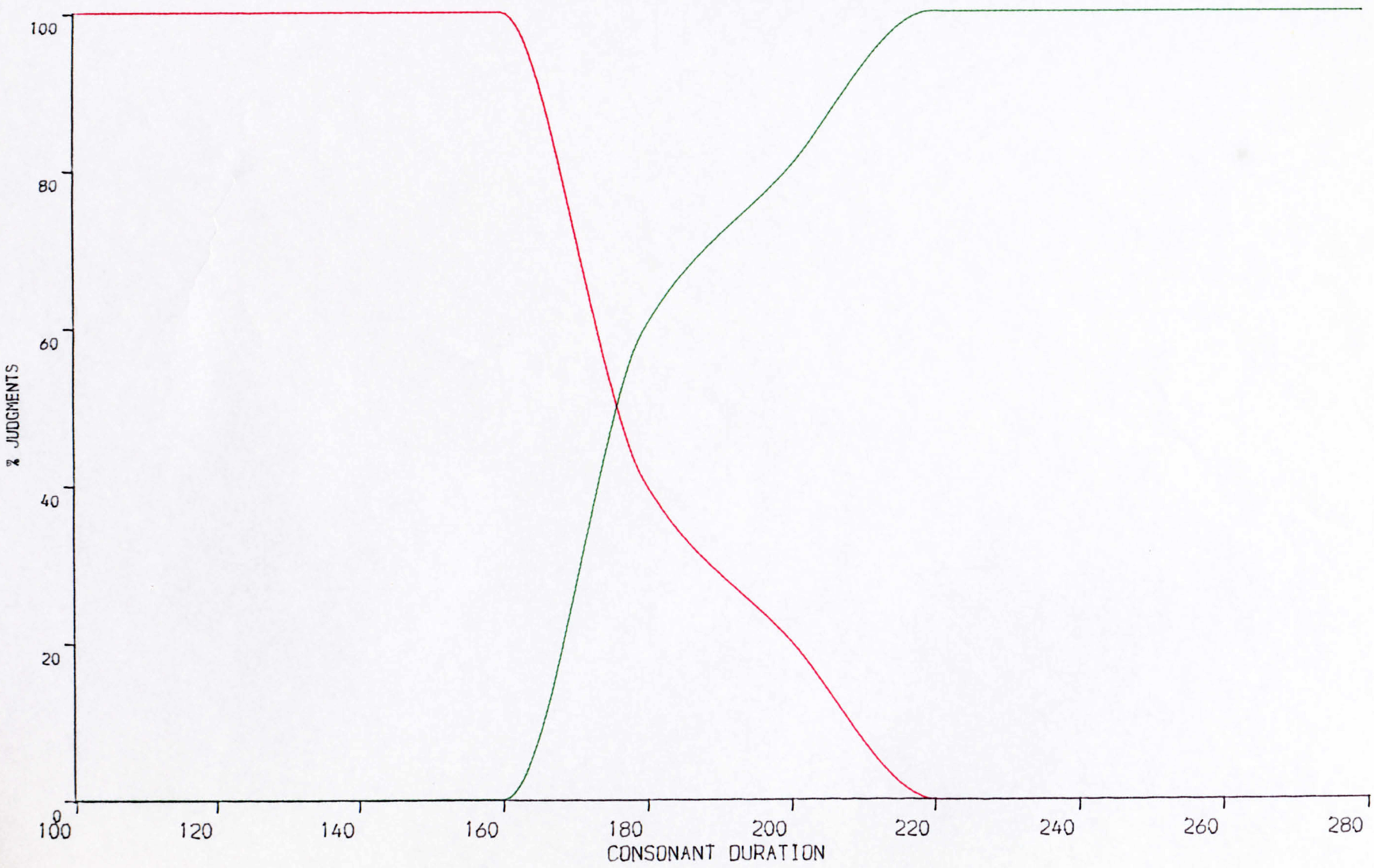


Fig. (4.18) /ħasan ~ ħassan/ Subject (AA)

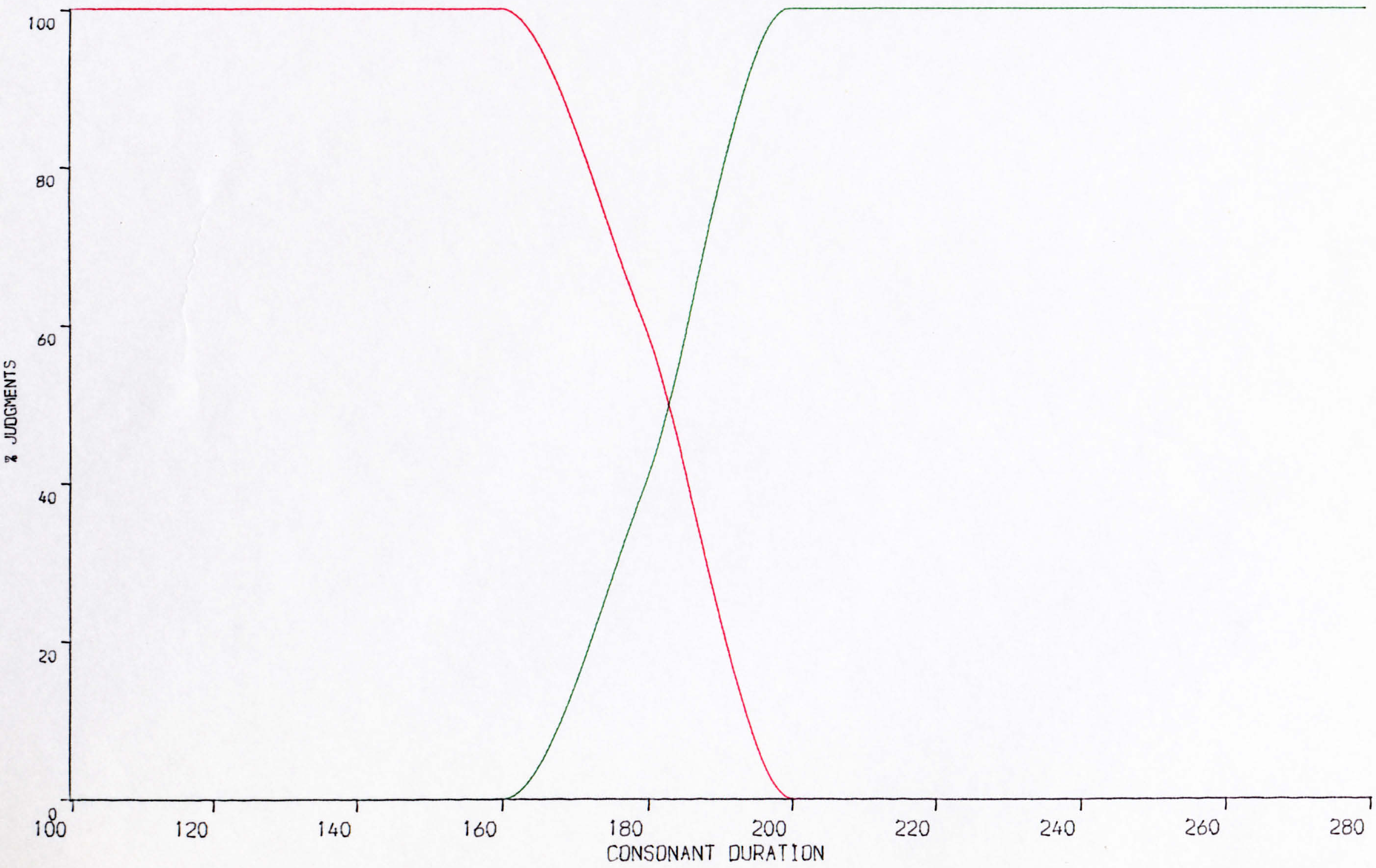


Fig. (4.19)

/badal ~ baddal/

Subject (AA)

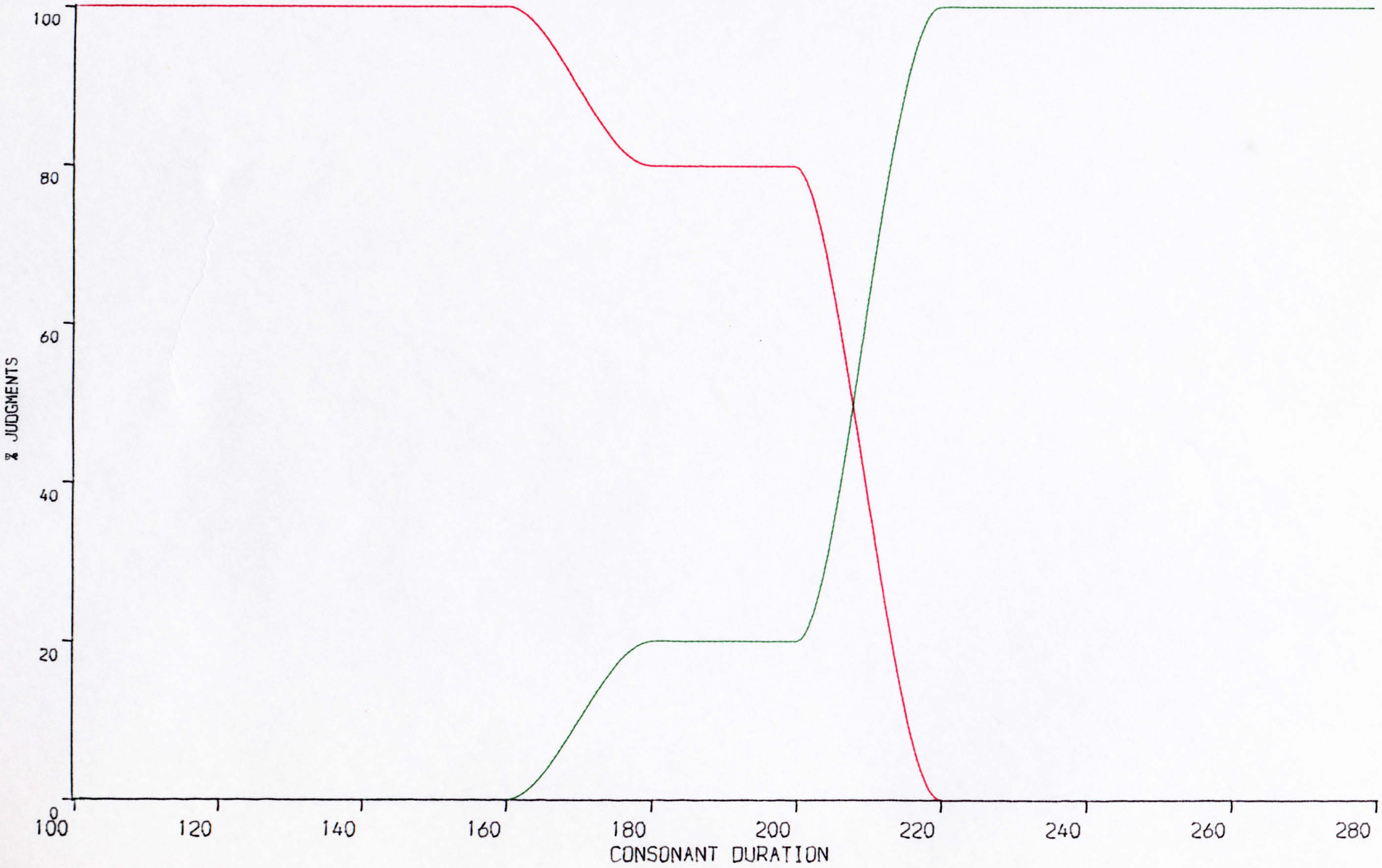


Fig. (4.20) /ħasan ~ ħassan/ Subject (IM)

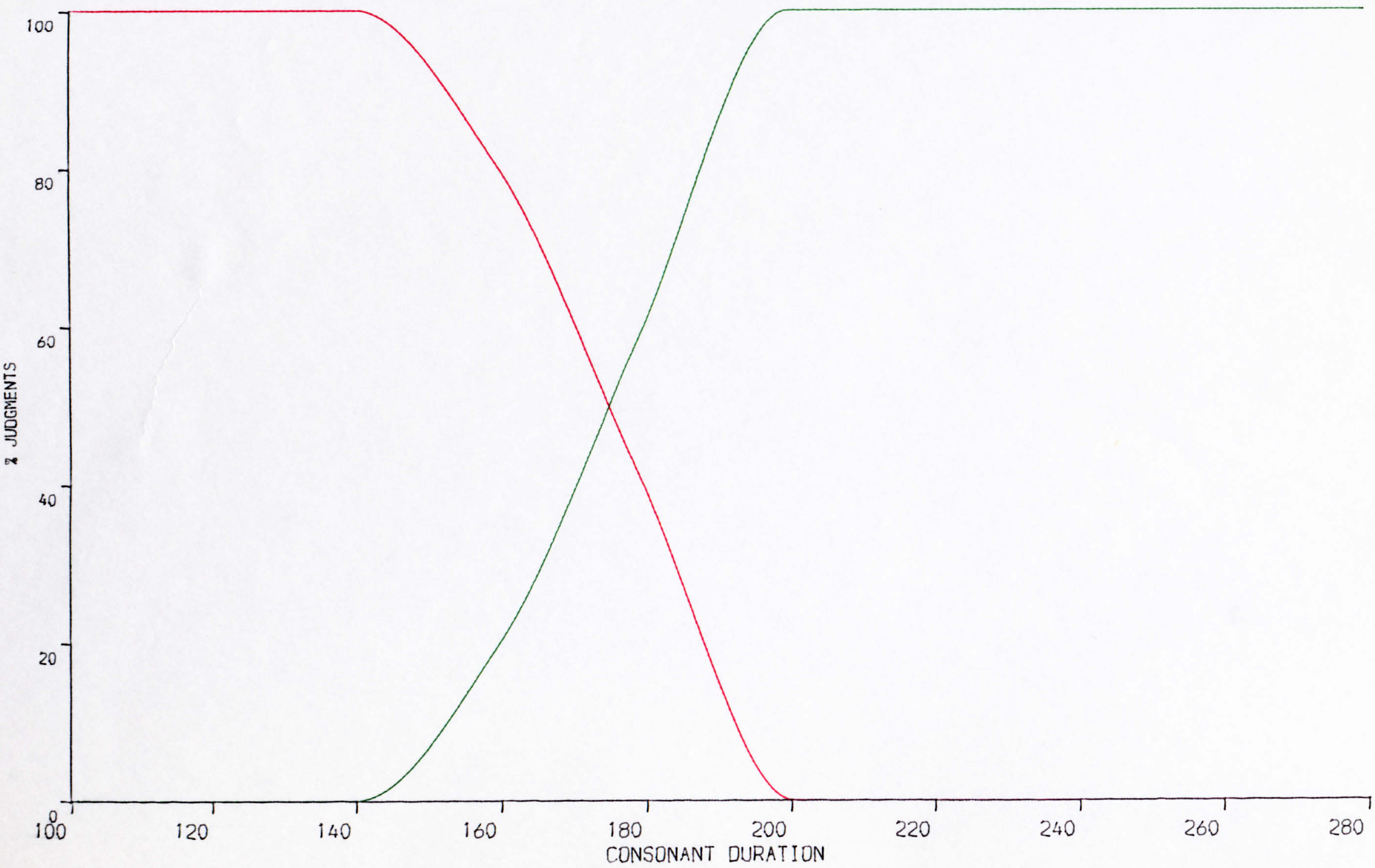


Fig. (4.21) /badal ~ baddal/ Subject (IM)

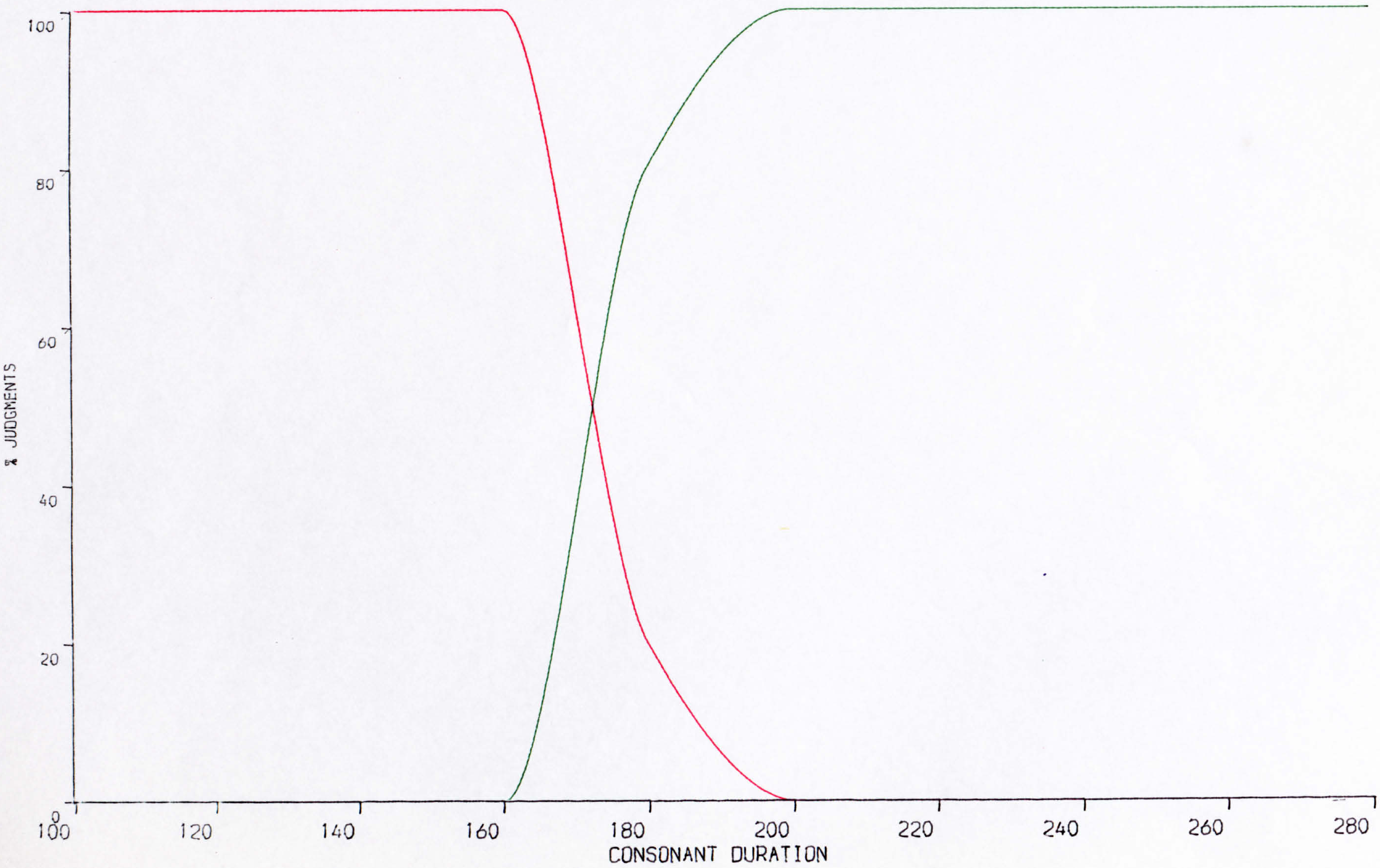


Fig. (4.22) /ħasan ~ ħassan/ Subject (AM)

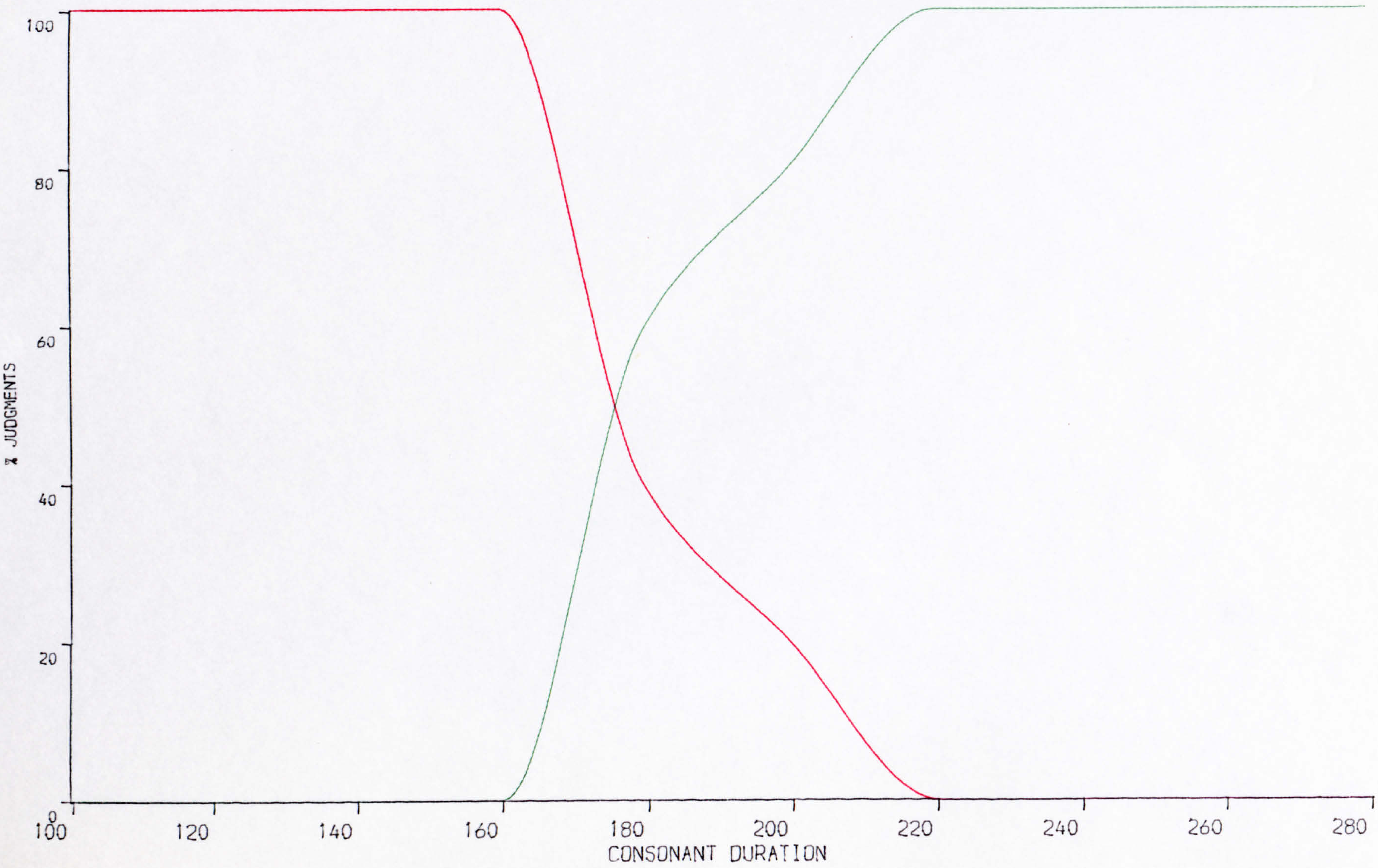


Fig. (4.23) /badal ~ baddal/ Subject (AM)

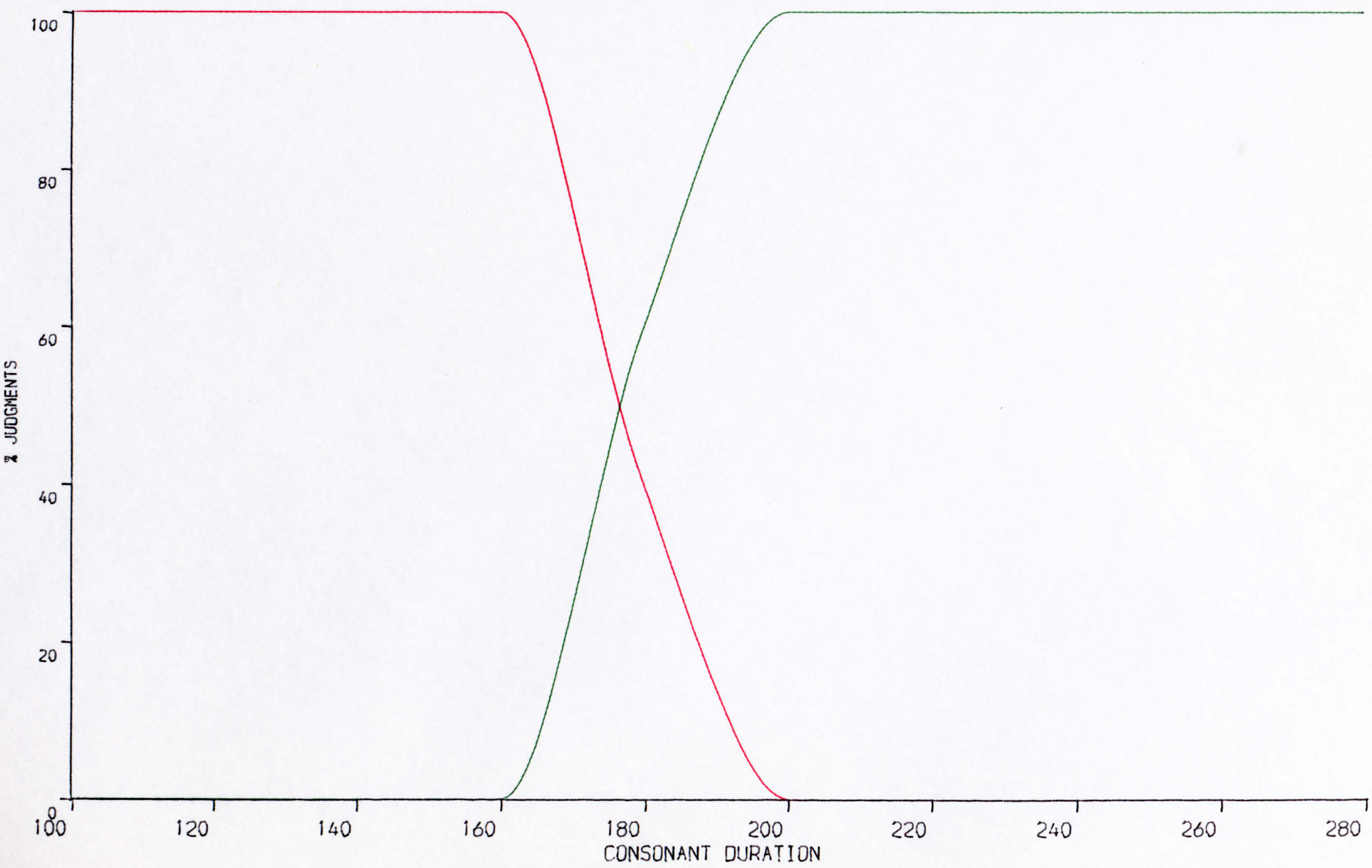


Fig. (4.24) /ħasan ~ ħassan/ Subject (TK)

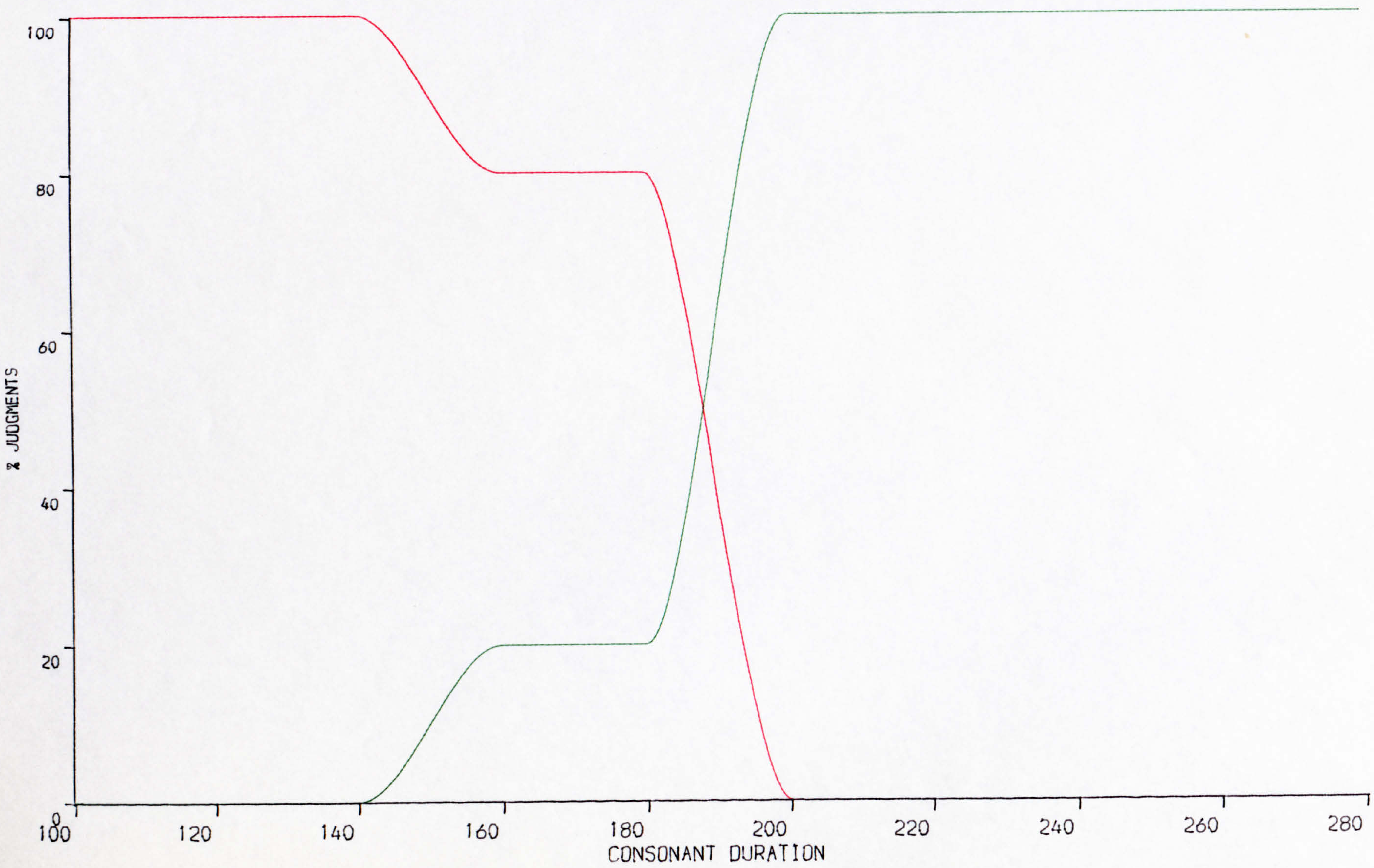


Fig. (4.25) /ħasan ~ ħassan/ Subject (AH)

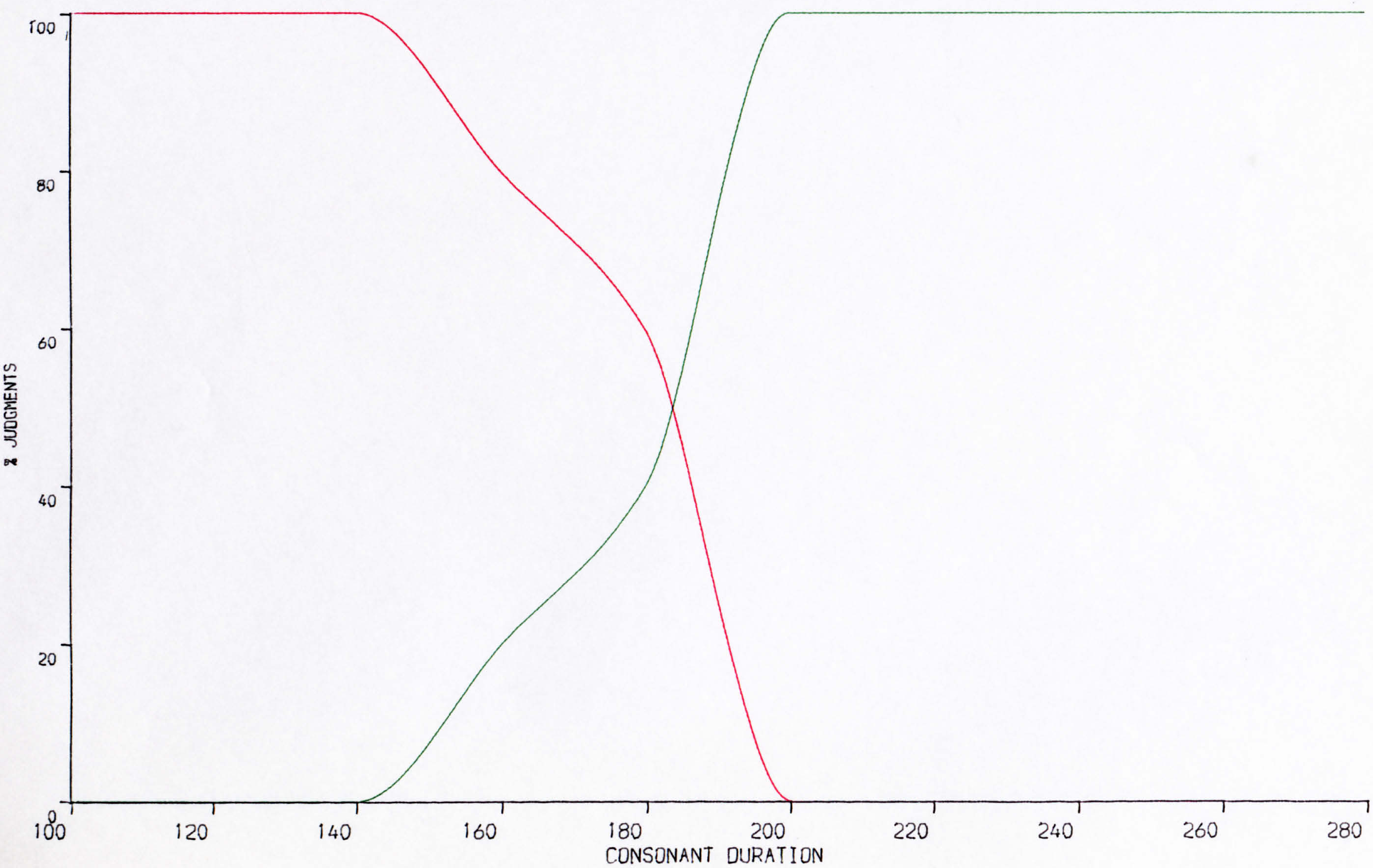


Fig. (4.26) /ħasan ~ ħassan/ Subject (FA)

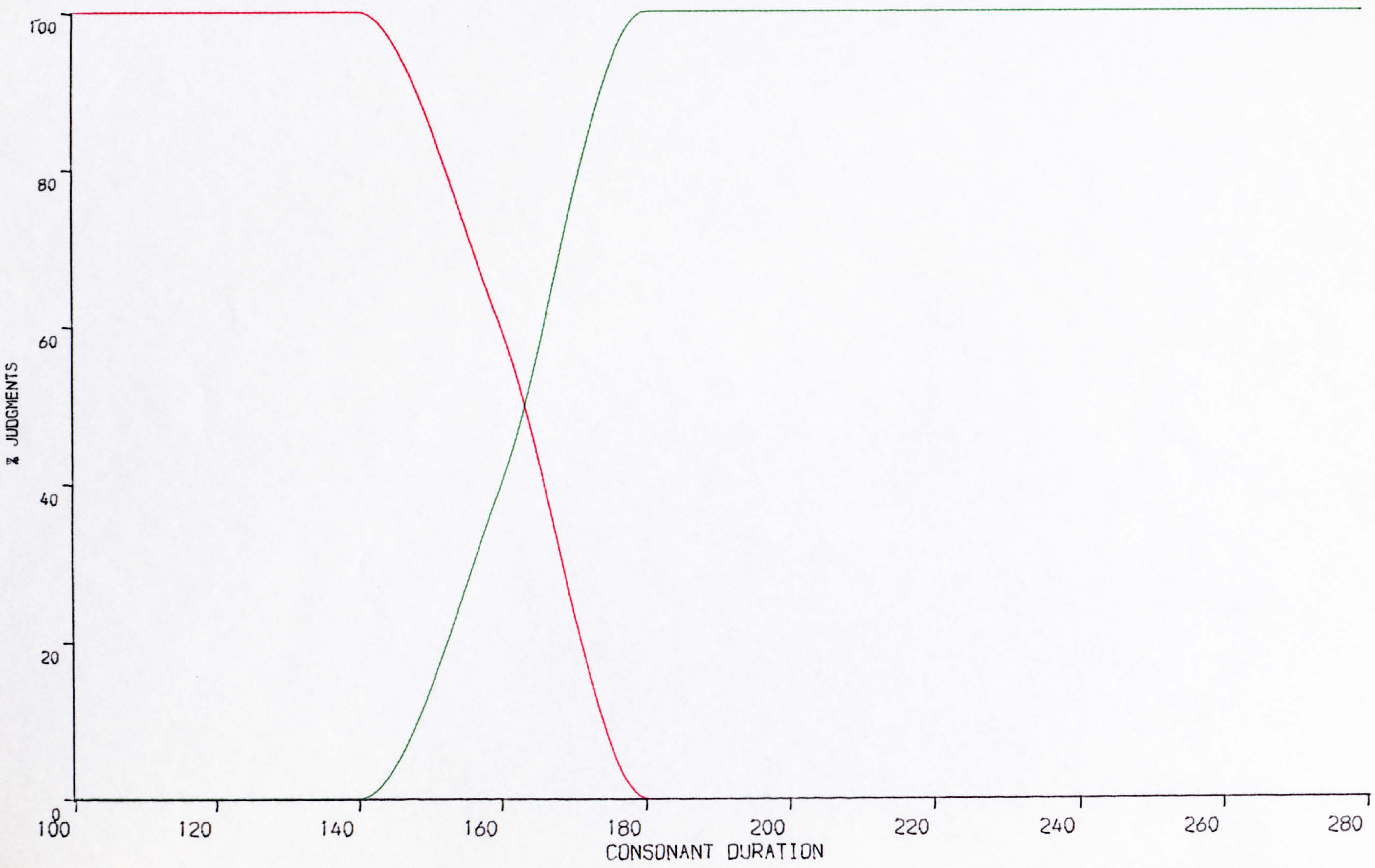


Fig. (4.27) /ħasan ~ ħassan/ Subject (HK)

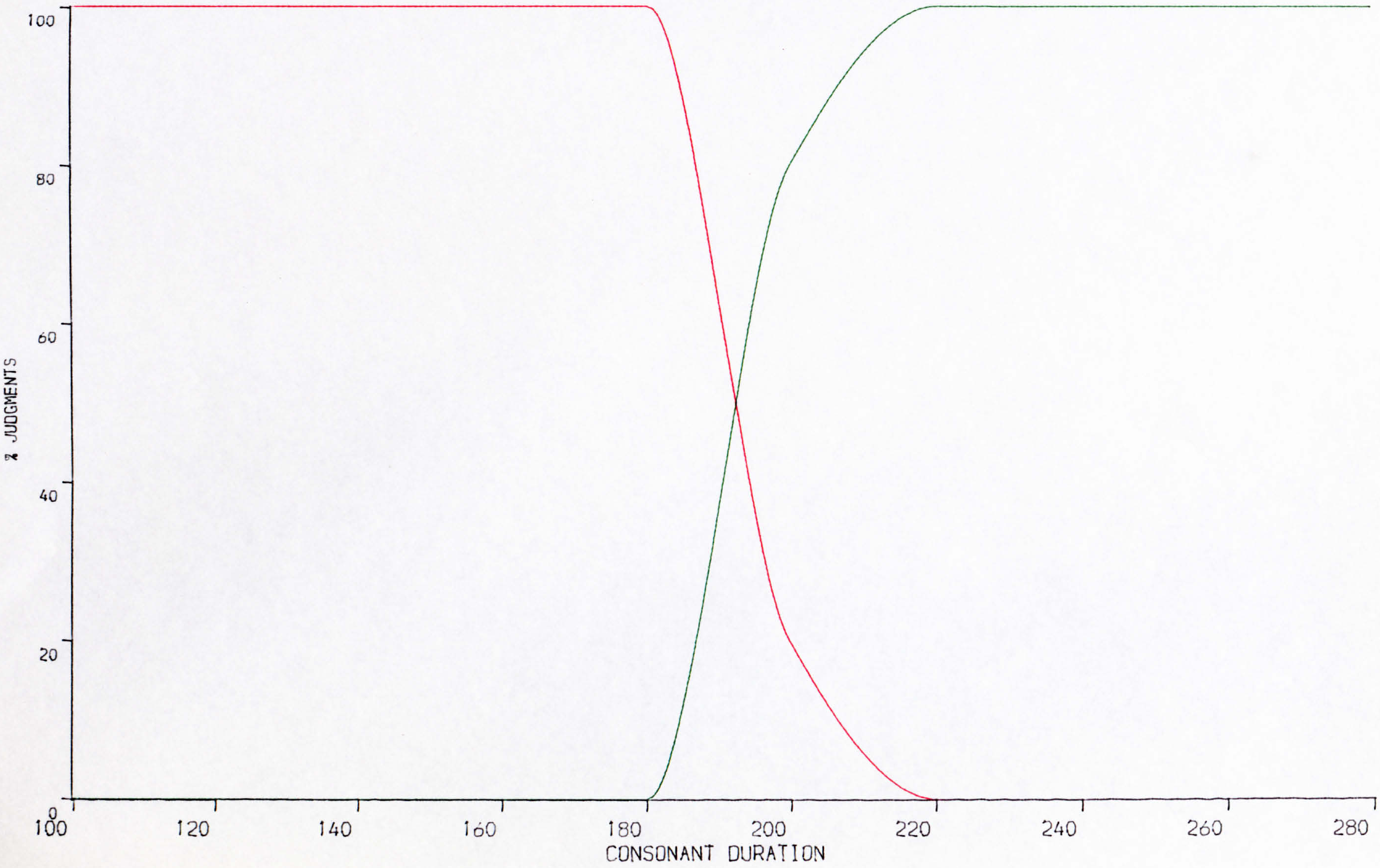


Fig. (4.28) /badal ~ baddal/ Subject (RK)

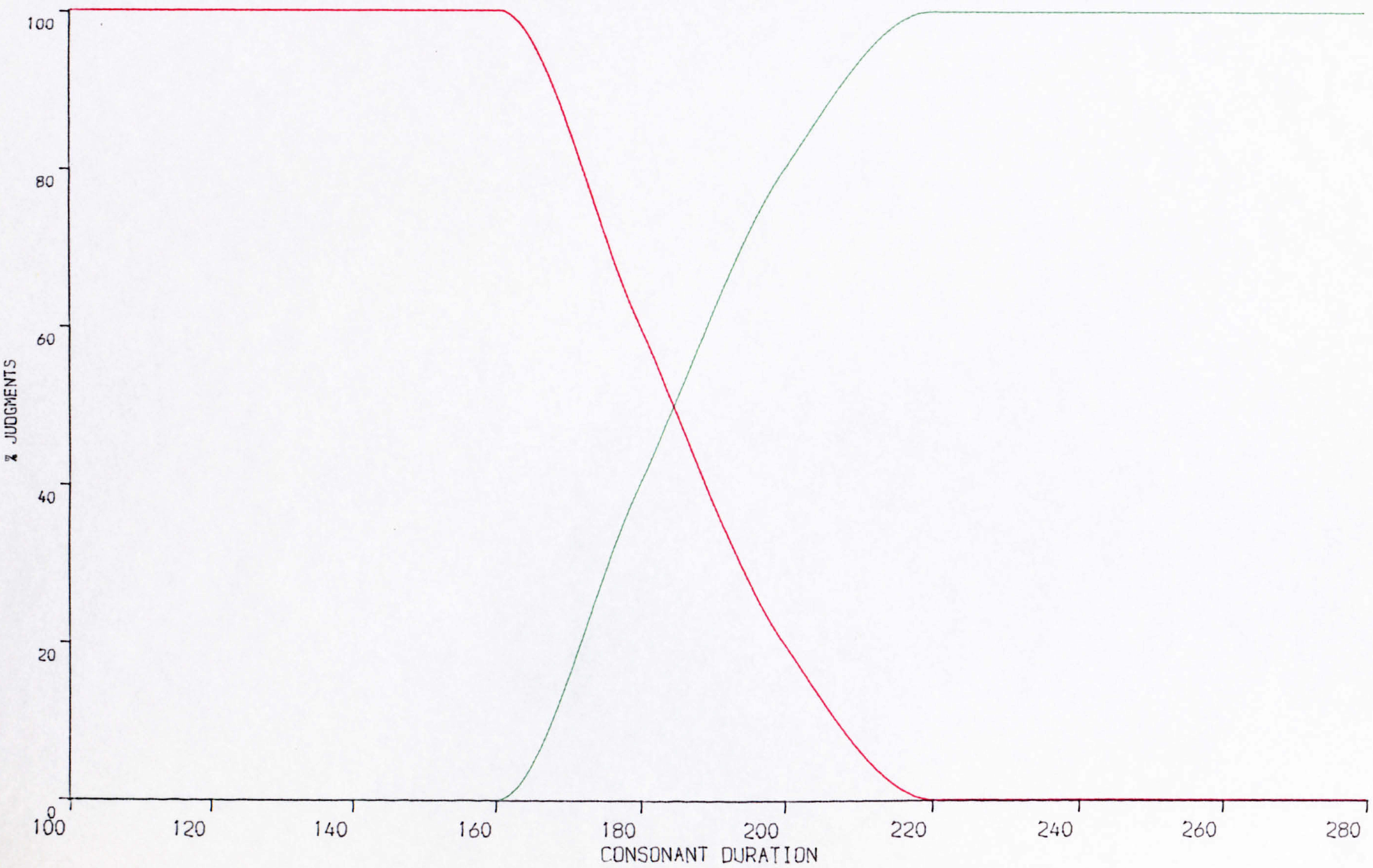


Fig. (4.29)

/badal ~ baddal/

Subject (MA)

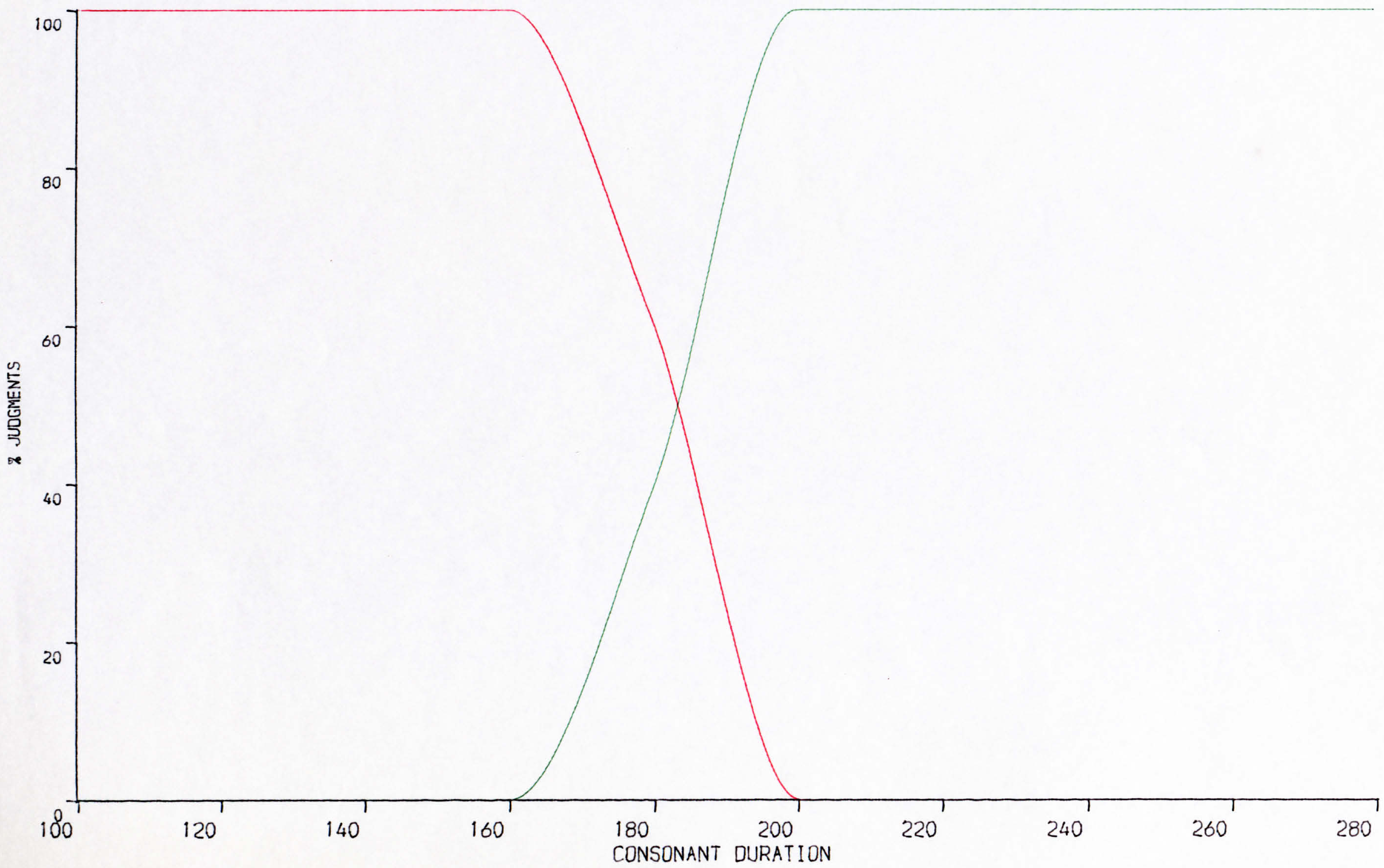


Fig. (4.30) /badal ~ baddal/ Subject (FN)

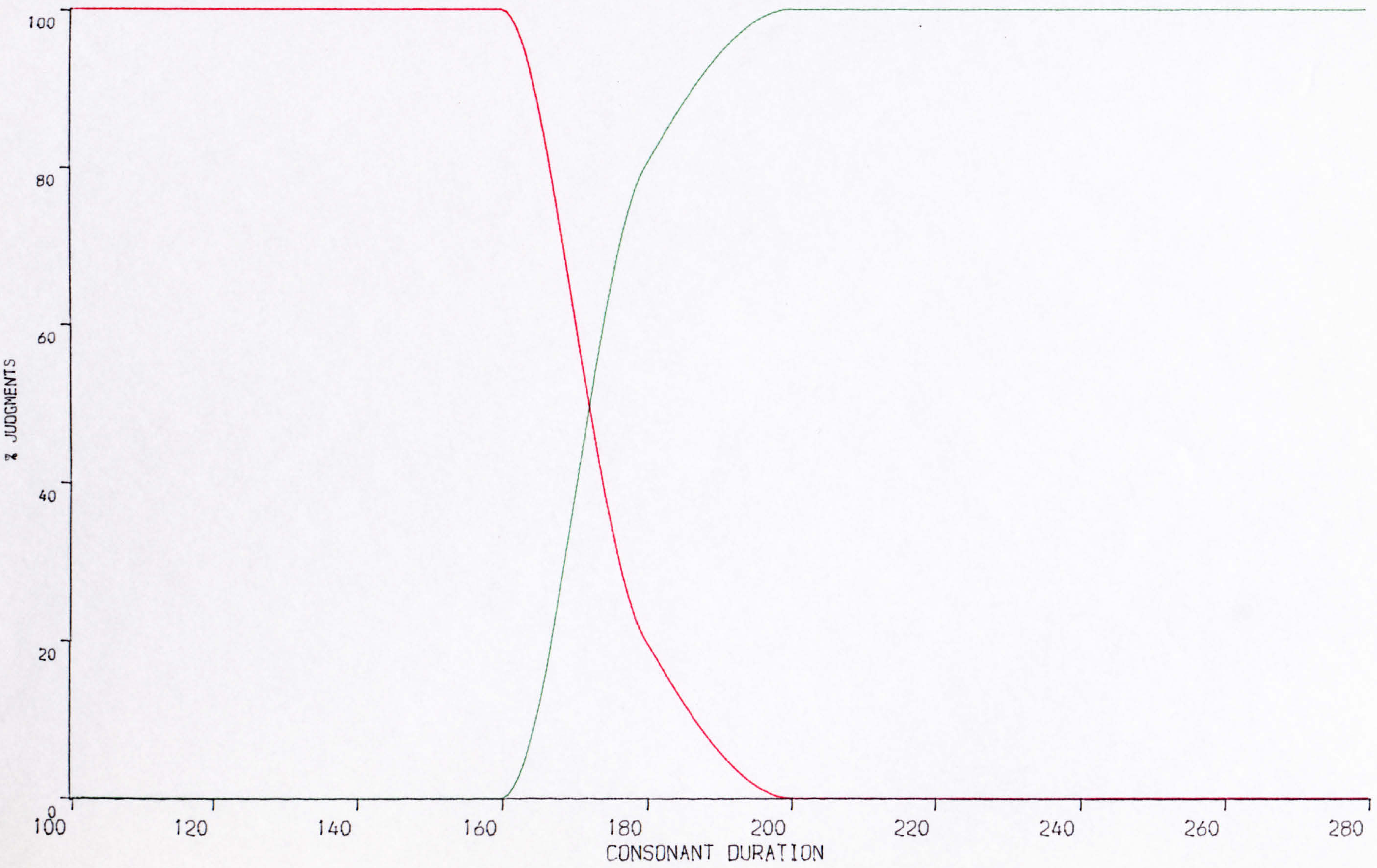


Fig. (4.31) /badal ~ baddal/ Subject (CG)

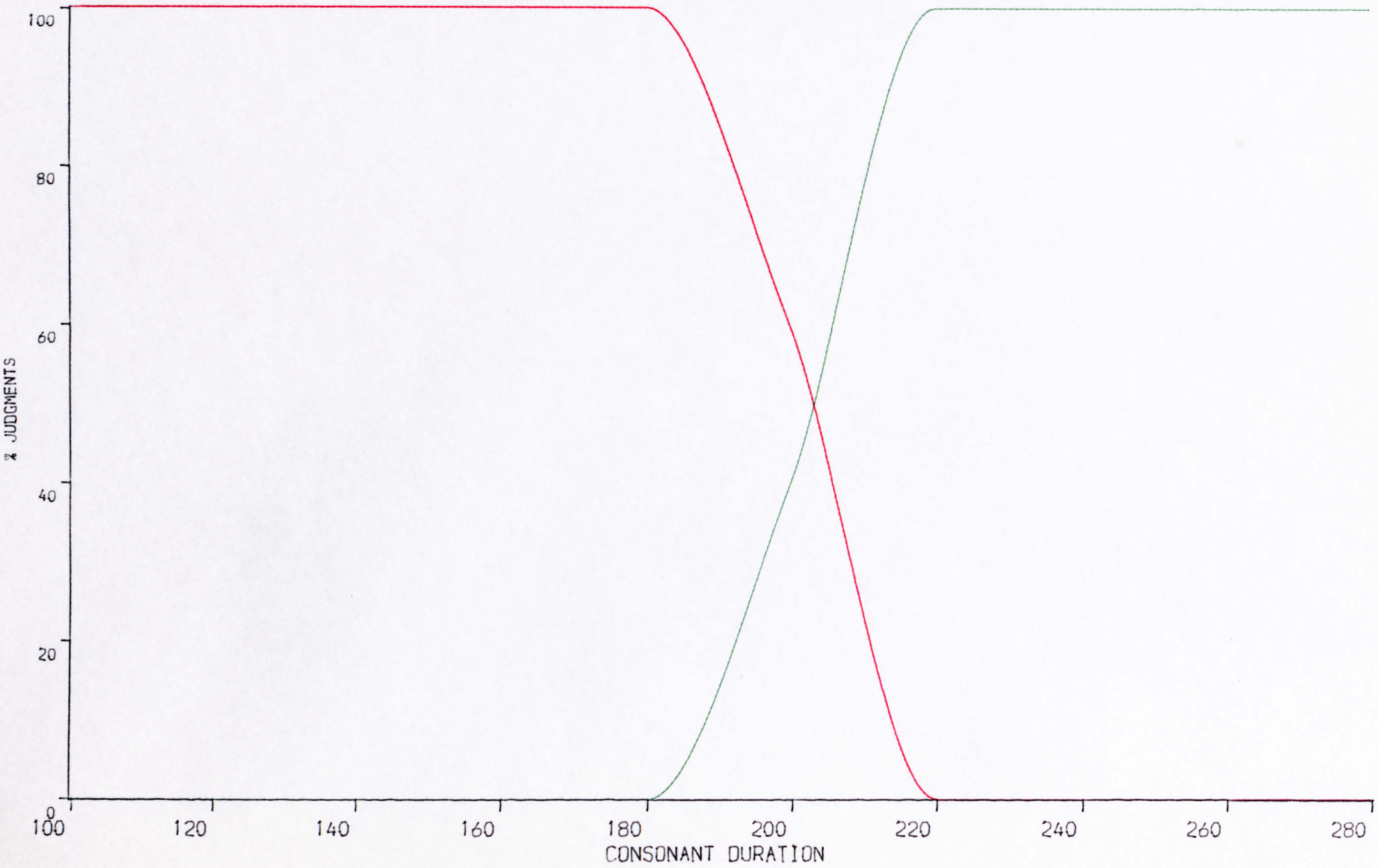


Fig. (4.32) /ħasan ~ ħassan/ 10 Subjects

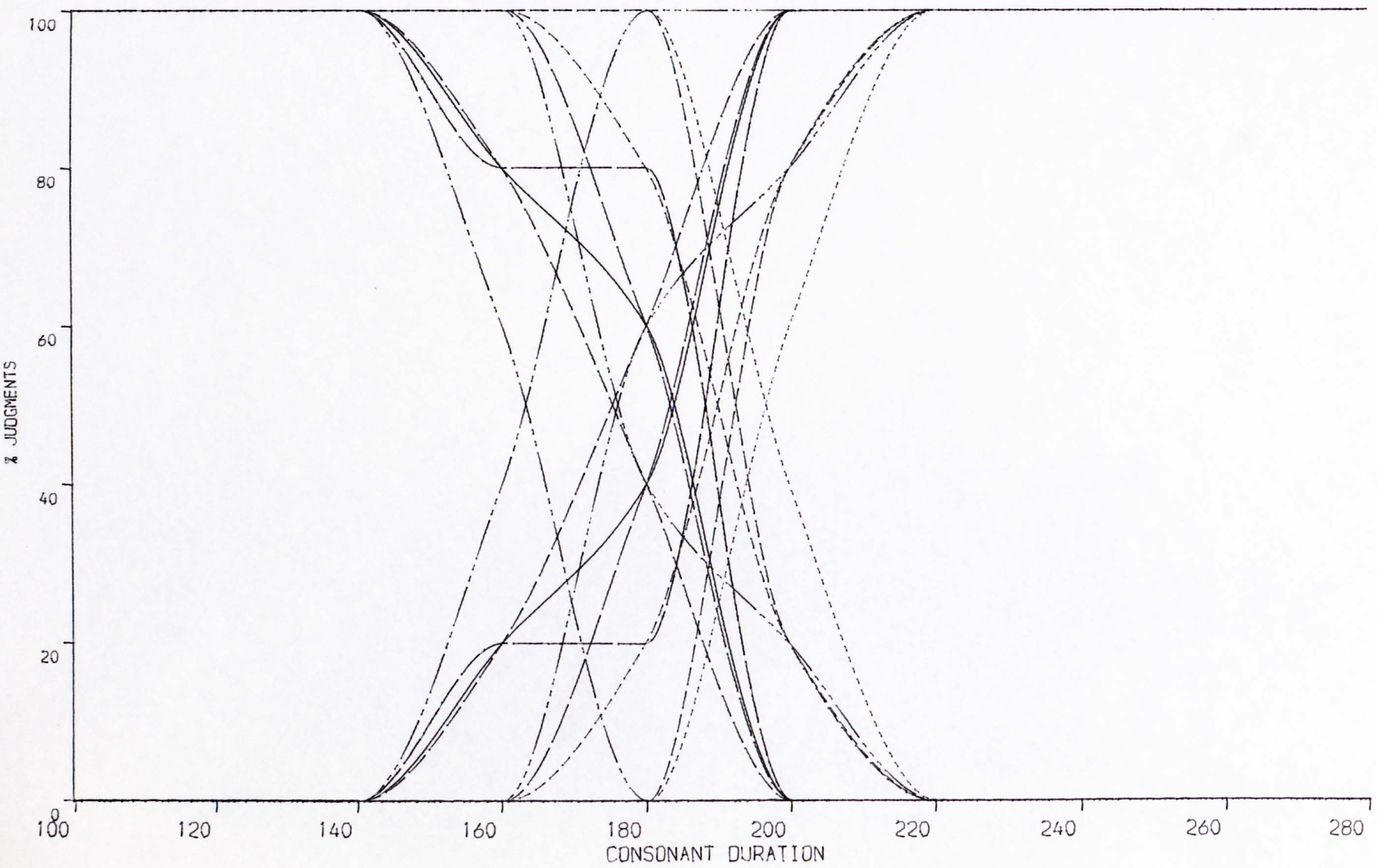
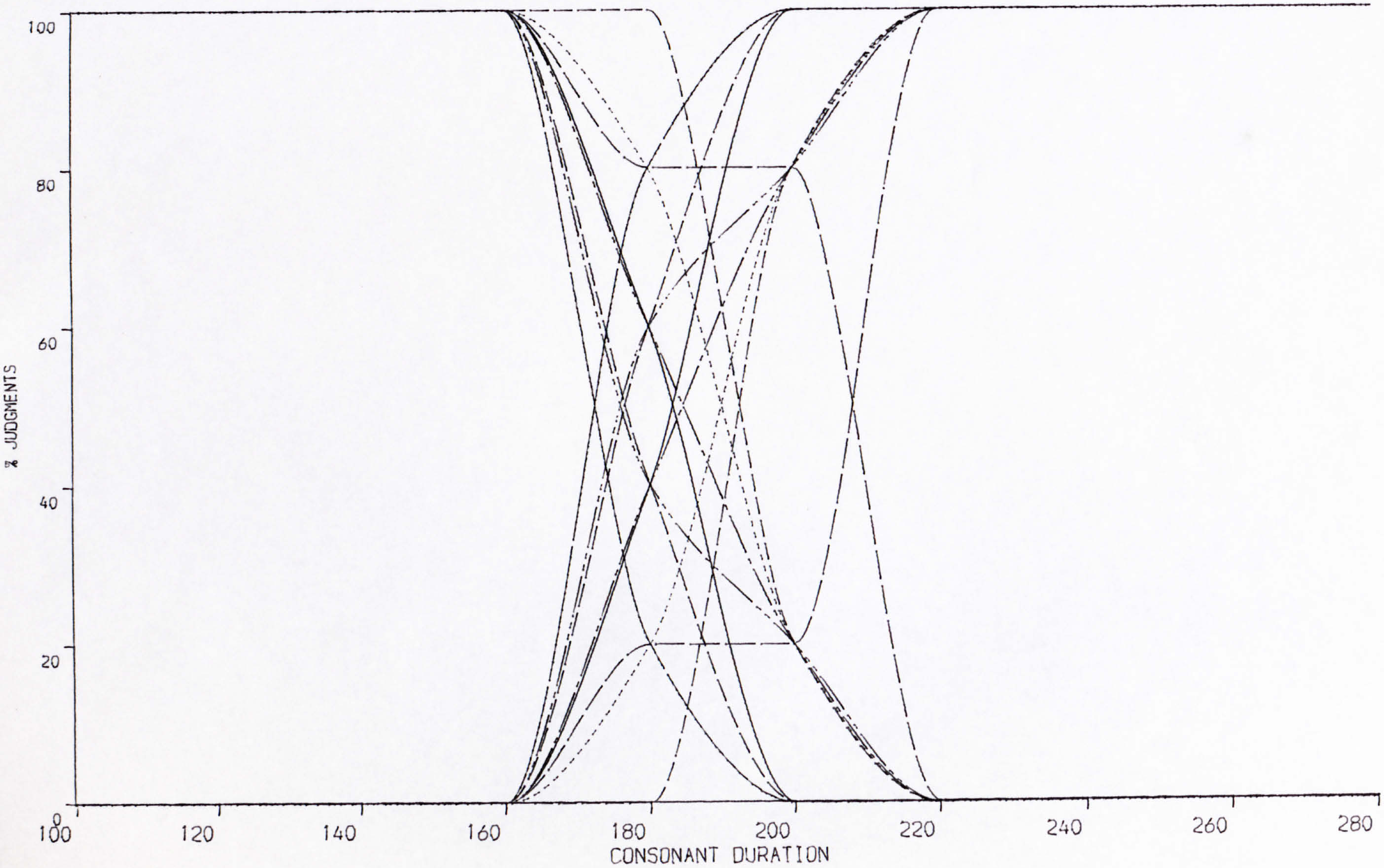


Fig. (4.33) /badal ~ baddal/ 10 Subjects



CHAPTER FIVE

PALATOGRAPHIC STUDY OF GEMINATE CONSONANTS IN I.C. ARABIC

5.1 Introduction and Aims

Palatography is generally defined as the technique used to register the location of contacts the tongue makes with the palate and teeth during speech. (Hardcastle, 1978). Generally speaking, two methods of palatography have been reported, namely the direct and indirect methods, which were originally developed by dental surgeons. (Coles, 1872; Kingsley, 1880). Since then, the two methods have been largely exploited for linguistic research by a number of scholars (e.g. Witting, 1953; Anthony, 1954; Abercrombie, 1957; Firth, 1957; Ladefoged, 1957; Jones, 1960; Strenger, 1968) among others. Both techniques are now included under the more general heading 'traditional or static palatography'. (Hardcastle, ibid.).

Abercrombie (1957, p.21) uses the term 'direct palatography' to mean the investigation of articulatory movements by means of marks made directly on the roof of the mouth, as distinct from the more usual technique which employs an artificial palate. In direct palatography, the palate of the subject's mouth is directly and completely coated with some substance; very often a mixture of powdered charcoal and chocolate. Palatographic records of the articulated sounds, normally known as palatograms, can be made for the subject's palate immediately after some of the substance has been wiped off by the tongue. The wipe-off represents the location of the tongue contact with the palate.

On the other hand, in the indirect technique of palatography, instead of covering the palate directly with the powdered mixture, an artificial palate, usually made of acrylic or vulcanite, is painted with some material which will be removed by the subject's tongue when contact is made. The artificial palate is then inserted into the subject's mouth and, after articulating the investigated sound, it is immediately withdrawn to be carefully inspected by the investigator at leisure. The wipe-off of the colouring medium on the artificial palate is taken to represent the regions of tongue contact with the palate.

Despite the fact that the indirect method is cleaner, more efficient and gives sharper outline of tongue contact than the direct method (Hardcastle, op.cit.), the latter method has a number of important advantages over the former one. For instance, some researchers (e.g. Abercrombie, 1957; Ladefoged, 1957; Hardcastle, 1971) remark that it is possible with direct palatography to study and compare the articulations of a relatively large number of subjects without a great deal of preparation; whereas indirect palatography entails the manufacture of an artificial palate for each subject which makes the whole process time consuming particularly as a plaster impression of the subject's palate must first be taken. Hardcastle (op.cit., p.100) reports that results from indirect palatographic investigations are thus usually of limited value as they are based on the study of a few individuals only. Unlike the indirect method, the direct method of palatography need not affect the normal articulation of a sound owing to the presence of the artificial palate as a foreign object inside the subject's mouth. It also enables the investigator to examine a far wider range of articulations due to the fact that the mixture powder employed in this technique can cover the soft palate and the teeth as well as the hard palate. It consequently permits the investigation of velar consonants and articulations involving contacts between the tongue and teeth. (Hardcastle, op.cit.).

In addition to the above mentioned advantages, the direct method has been extensively used for pedagogical purposes in many experimental phonetics laboratories because of its general simplicity and relative inexpensiveness. In this respect, Abercrombie (op.cit.) states that "Articulatory movements which are not visible are not easy to grasp for beginners in phonetics, whose kinesthetic sense is usually still undeveloped for the tongue. Descriptions of these movements can be made more real and less purely theoretical by means of palatography (which must be active participation in palatography, not just the exhibition of other people's palatograms); but the labour and expense involved in making artificial palates for each member of a large class would normally rule its use out." [p.25].

Although palatograms are normally made for single isolated sounds which involve the tongue-palate contact, it is also possible to produce a palatographic record of the entire word consisting of a composite picture made up of a number of continually changing contacts. (Hardcastle, 1978). The technique of word-palatography was extensively used to examine certain prosodic features in different languages by Firth (1948) and some of his disciples. Ladefoged (1957) points out that no matter which system of palatography is employed, i.e. either the direct or the indirect method, the results of a palatographic investigation will be articulatory data in the form of diagrams or photographs. He also remarks that any diagram or photograph of the palate is a two dimensional representation of information which was originally in three dimensions. Accordingly, palatograms often fail to convey important information concerning the shape and depth of the palatal cavity, and the position and slope of the alveolar ridge. Hardcastle (1971) comments saying that "Without such information, it is extremely difficult to compare the articulations of different speakers." [p.101].

Ladefoged (ibid., p.766) suggests what he calls the 'contour system' for providing a representation of the depth of the palate. According to this system, each palatogram is accompanied by a diagram showing a sagittal section of the palate. This can be obtained by sawing along the mid-line of a cast of the roof of the mouth. He claims that such a diagram, representing the sagittal section of the palate, will be inaccurate in so far as it fails to take into account the movements of the soft palate, which may be important in speech. But it will give a satisfactory outline of many of the important features of the roof of the mouth. He also contends that knowing the shape of the roof of the mouth, it is often possible to deduce the probable shape of the tongue during the articulation of the consonant in an utterance. After gaining experience with palatography, the information necessary for making a very approximate diagram of the presumed position of the tongue may be obtained by comparing a palatogram with a cast of the roof of the mouth.

However, despite the extensive use of the technique of static palatography, both the direct and the indirect method, in linguistic research, it has been found to be limited in the investigation of speech sounds in that it gives a single static representation of what is essentially a dynamic event. (Hardcastle, 1978). Furthermore, this technique primarily involves the insertion or application of some foreign body or substance, viz. an artificial palate or a powder mixture, into the subject's mouth which may possibly, though minimally, cause distortion of normal articulation.

Speech researchers (e.g. Kydd and Belt, 1964) believed that in the past there had been no reliable method of sequentially determining where the tongue comes in contact with the hard palate during connected speech. Cinefluorography, for example, has been used to provide some valuable spatio-temporal data on articulatory activities of the speech organs and to demonstrate that the tongue is in contact with the palate. But this technique gives no indication as to the area of tongue contact and it does not show the lateral areas of the tongue in contact with the palate and teeth. What is more important, it is severely limited by the risk of radiation damage. (Hardcastle, 1970, 1972). Other methods, such as magnetometer, ultrasonic techniques and jaw movement transducers are being developed as alternatives to X-ray but they do not show tongue-palate contact area either. On the other hand, despite being useful techniques in studying articulatory kinematics, spectrographic measurements and electromyographic recordings provide only indirect evidence of actual movements of the articulators which must be inferred either from electrical activity in the muscles, which precedes them, or from the acoustic vibrations in the air, which succeed them. (Butcher and Weiher, 1976). Not only does electromyography precede and acoustics follow but also (a) when electromyography is done, usually only a small fraction of the possible forces are monitored, so the picture conveyed may be misleading, and (b) whereas a given vocal tract shape defines precisely its acoustic properties as a resonator, the reverse is not true: a given set of resonances (formants) can arise from different vocal tract shapes. That is, the mapping from acoustics back to articulation is likely to be one-to-many, not one-to-one. Also, the relationships between articulatory movements and acoustic pattern changes are non-linear, i.e. in some cases a small articulatory change will result in very little acoustic change, whereas in other cases, from the same small articulatory change, a large acoustic effect may result.

However useful in speech research electromyography is, in so far as the tongue is concerned, it is limited largely to investigating electrical potentials associated with contraction of some of the extrinsic lingual muscles only; besides it "has to be largely limited to the use of surface electrodes, rather than of needle or hooked-wire electrodes, thus making the interpretation of data difficult." (Hardcastle, 1970, p.54). More recently, however, bipolar hooked-wire electrodes have been used by some researchers (e.g. Tuller et al. 1981a, 1981b) to record electromyographic activity from the anterior portion of the genioglossus muscle.

A device designed to delineate the continuous movement of the tongue temporally and spatially would be of great benefit to speech scientists for obtaining detailed information on the dynamics of lingual articulation. The need of such devices has greatly encouraged the development of electropalatography; a technique which aims primarily to provide information about one particular aspect of lingual activity, namely details of the exact location and timing of contacts that the tongue makes with the palate during connected speech. (Hardcastle, 1972). It is only by means of an artificial palate in which is embedded a number of electrodes, that the temporal and spatial ordering of the points of contact between tongue and palate can be accurately recorded. There has been an interest in recording these contacts for a long time, but it is only in recent years that successful techniques have been developed to give displays of tongue palate contact as a function of time. Hardcastle (1968) states that "Most of these systems[1] use some sort of artificial palate incorporating numbers of sensing elements, which are activated when touched by the tongue." [p.53].

Apart from its obvious use in experimental phonetic research, there are a number of practical applications of electropalatography, such as in teaching foreign languages and in helping deaf children acquire speech. This technique, however, has its own limitations as well. For example, there is a possibility that the artificial palate may interfere with normal articulatory patterns, although the interference from such a foreign body with articulation can be minimized by making the artificial palate thin enough. Moreover, electropalatography is only suitable for investigating articulations involving lingual contact with the palate and close or half-close front vowels. It gives no information as to articulations where no tongue-hard palate contact is made, such as those sounds involving contact in the velar region and further back, and some open vowels. There also exists the possibility of the danger of forgetting that when no contact shows, there is no significant tongue-palate contact at all. Besides, the technique suffers from the difficulty of designing a 'universal' artificial palate basically because of the wide differences in palate shapes

---

1. The various different systems of electropalatography have been fully described by Hardcastle (1970, 1971, 1972). For more recent works dealing with the use of this technique in linguistic research, see articles printed in 'Work in Progress' vol. 1 (1977), vol. 2 (1978) and vol. 3 (1981), issued by the Department of Linguistic Science, University of Reading. Also see articles printed in the Annual Bulletin, Research Institute of Logopedics and Phoniatrics, Faculty of Medicine, University of Tokyo, No. 6 (1971), No. 12 (1978) and No. 17 (1983).

and sizes among individuals. It is hoped that this difficulty will be overcome by the possibility of making a universal palate consisting of strips of malleable material which could be stuck onto the subject's palate with dental adhesive. (Hardcastle, 1972, p.211).

This brief introduction to the technique of palatography leads us to our present experiment which was designed to stand by itself to give further support to our preceding as well as to our succeeding experiments. It was limited in scope; the direct method of palatography and electropalatography (EPG) as a pilot attempt were employed. With the aid of these techniques, we aimed to test our primary hypothesis to see whether or not the contact between tongue and palate is firmer and greater in area for word-initial and word-medial geminate consonants than for their non-geminate partners in I.C. Arabic.

5.2 Experiment No. 1 (Direct Palatography)

5.2.1 Experimental Procedure

5.2.1.1 Selection of Stimuli

The test items of this experiment were not the same as those studied in our other experiments. A new list was prepared and the new stimuli were carefully chosen to ensure that only one segment in each word was capable of producing a wipe-off from the marking substance with which the palate was covered before the tongue-palate contact was executed. The chosen words represent two oppositions of different contoid types where contrasts between single [s] and [d] and their geminate cognates [ss] and [dd] were involved in word-initial and word-medial positions.

The new list comprised 16 words arranged in minimal pairs. All words were bisyllabic bearing a primary stress on the first syllable. The second syllable was either of the pattern (-CV) or of the pattern (-CVC), where the V-element was either [i], [a] or [u]. The following is the complete list of the words under investigation:

Words with single consonants	Meaning	Words with geminate consonants	Meaning
/ξada/	except	/ξadda/	he counted; he let pass
/ξasa/	let's hope	/ξassa/	he inserted
/daaξi/	requisite; claimant	/ddaaξi/	the requisite; the claimant
/saaξi/	messenger	/ssaaξi/	the messenger
/ħasab/	noble birth	/ħassab/	he reckoned
/hadaf/	aim	/haddaf/	he aimed (at sth)
/suħub/	clouds	/ssuħub/	the clouds
/digam/	buttons; switches	/ddigam/	the buttons; the switches

5.2.1.2 Subjects

The author of the present study was his own subject in this experiment and, therefore, the palatograms obtained represent photographic records of his own palate. He has essentially normal oral and dental structures as well as normal speech and hearing.[1]

5.2.1.3 Instrumental Set-up and Method

For the purpose of obtaining palatograms for the words under investigation, the experimenter's palate was first thoroughly sprayed with a mixture of powdered charcoal and drinking chocolate. The mixture was composed of 2/3 charcoal and 1/3 chocolate. Each word was then uttered five times in succession, and as similarly as possible on all five occasions, to ensure a satisfactory amount of wipe-off for the photograph. While being pronounced, the speech was simultaneously recorded on high quality tape (type Ampex 1200) so that the recording might be played back when examining the resulting palatograms at a later stage. The tape-recorder used (type Ferrograph Logic 7) was placed on a table in front of the experimenter.

After producing each of the investigated words, and without swallowing or moving the tongue, the experimenter placed his wide-open mouth over a plane mirror that stood opposite to a 35-mm camera (type Exakta, model VX1000). The palatal reflected pictures were then graphically recorded by means of the camera which was fitted with a standard Carl Zeiss-Tessar lens. The lens was, in turn, attached to a convertor for close-up photography and extra magnification of image. The camera was supplied with a black-and-white film (type Ilford, FP4, 125ASA). The aperture was set at f/11 and the shutter-speed at 1/60 sec. Care was taken to

---

[1] See section 3.2.2 for further information about his linguistic background.

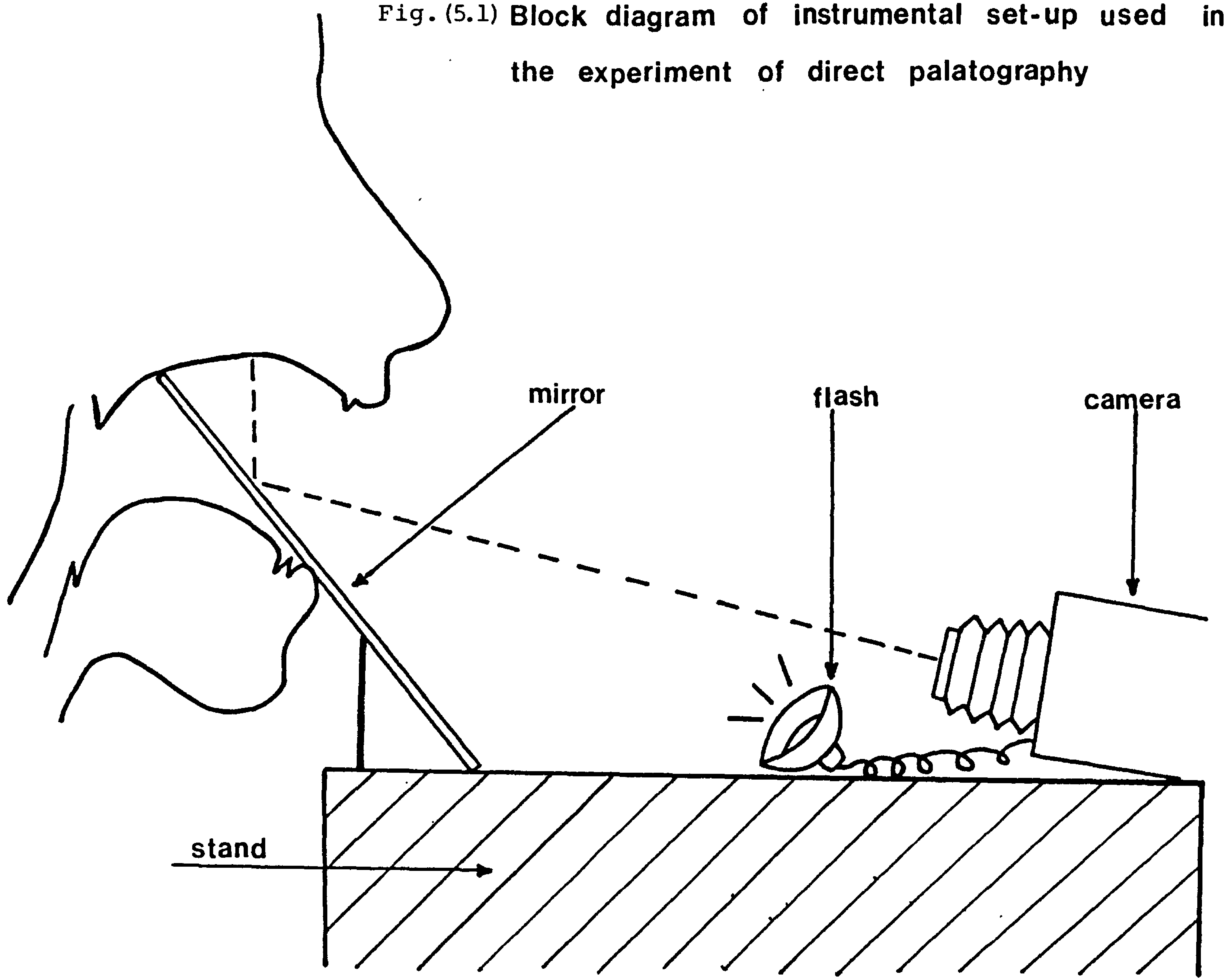
ensure that the experimenter did not exhale and fog the mirror which could distort or blur the palatal image reflected on the mirror. Therefore, the mirror was repeatedly cleaned with a soft piece of cotton cloth.

In order to have sufficient source of illumination a flash-unit (type Vivitar 45) was positioned at a suitable distance opposite to the experimenter's mouth. The flash-unit was linked to the camera via a synchronization cord for off-camera flash. For each investigated word, two photographic frames (i.e. exposures) were taken so that in the event of the failure of one frame the other should be used. After each item was graphically recorded, the experimeter rinsed out his mouth thoroughly, thus leaving both tongue and palate free of the mixture powder. Diagram 5.1 illustrates the instrumentation employed in this experiment.

5.3 Segmentation Criteria of Palatograms

The obtained palatograms were carefully inspected. They were divided into zones similar to those suggested by Firth (1948). Firth presented articulatory terms as a first attempt at zoning the palate on the basis of the dentition plan. To these zones he also added a natural median line, and right and left lines parallel to the median starting from the interstices between the front and lateral incisors. According to him a perfect palate may look like the one illustrated in Fig. 5.2. The reference lines and the zones shown in this figure may perhaps ensure an adequate representation of the significant features of the palatograms. Each zone is forward of the line opposite which it is entered

Fig. (5.1) **Block diagram of instrumental set-up used in the experiment of direct palatography**



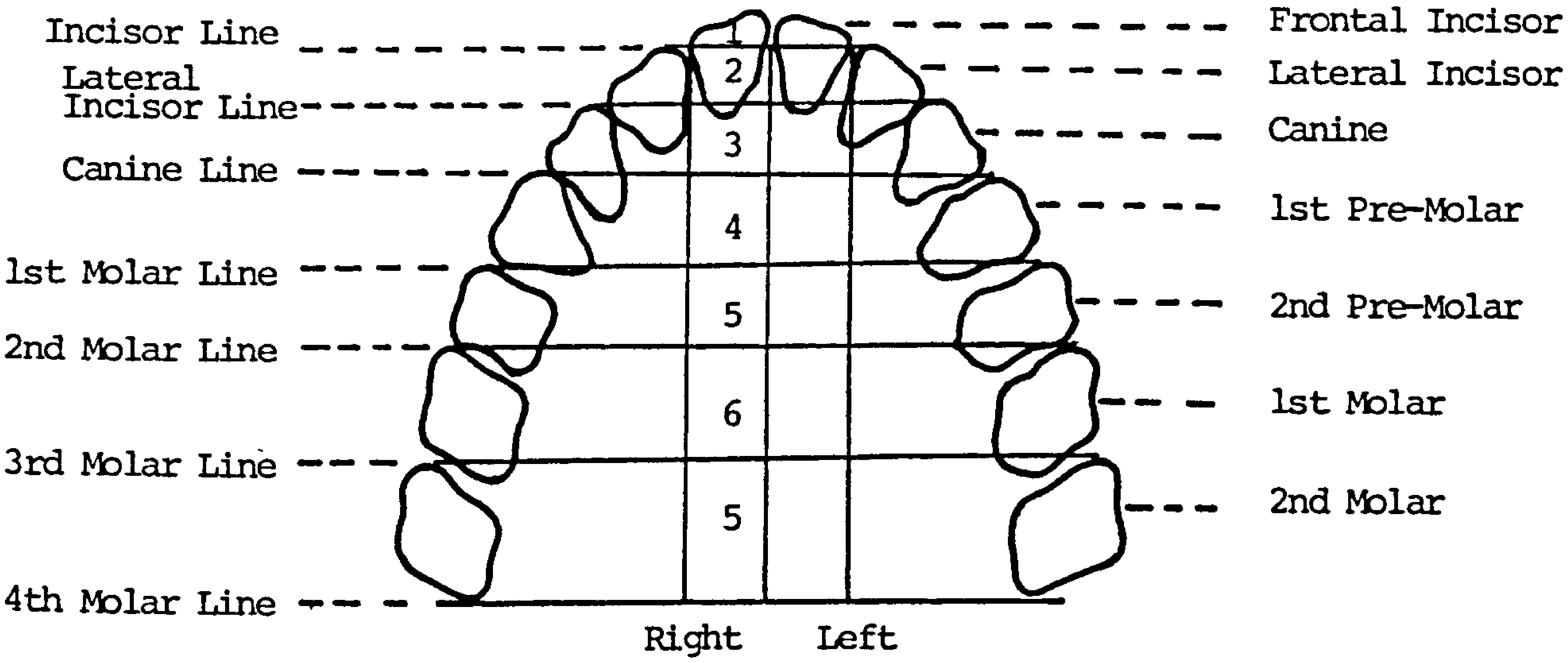


Fig. (5.2) Schematic diagram showing the palatal zones and lines as suggested by Firth (1948).

	The horizontal lines	The zones	The grouped zones
1	Incisor line	Dental	Dental
2	Lateral incisor line	Denti-alveolar	
3	Canine line	Alveolar	Alveolar
4	First molar line	Post-alveolar	
5	Second molar line	Pre-palatal	Palatal
6	Third molar line	Mid-palatal	
7	Fourth molar line	Post-palatal	

The vertical lines	The vertical zones
The median line	Right and left zones : to speaker's right and left of the median line
The right line	Right alveolar zone : to the right of the right line
The left line	Left alveolar zone : to the left of the left line

Table (5.1) The terminology used by Firth (1948) to describe the various palatal lines and zones.

in table 5.1. The two words 'right' and 'left' refer to the speaker wearing the artificial palate.

The following palatal zones were considered appropriate for our present experiment:

1) dental 2) denti-alveolar 3) alveolar
4) post-alveolar 5) pre-palatal 6) palatal
7) post-palatal.

Another eighth zone can be added to the above seven zones. This zone may be termed 'velar'. By definition the velar zone is taken to represent a fifth molar line added to those horizontal lines described by Firth (op.cit.). Since the velar zone was not quite obvious in most of our palatograms, therefore, it was not included in our classification of the palatal zones.

5.4 Analysis of Palatograms

5.4.1 Palatograms illustrating articulations of word-medial consonants

Palatograms 1-4 represent the words /ξasa/, /ξassa/, /ħasab/ and /ħassab/, respectively. They show the articulations of single and geminate fricative consonants occurring word-medially. In all four palatograms it is obvious that the marking medium is removed from zones 3 and 4, i.e. in the alveolar and post-alveolar regions, as well as along the sides in all zones. In palatograms 2 and 4 there is a clearer and more marked wipe-off than that in palatograms 1 and 3. This suggests that the tip or blade of the tongue touches the alveolar and post-alveolar regions more firmly during the articulation of the geminate [ss] than during the articulation of the non-geminate [s] in word-medial position. It is noticeable in palatograms 1 and 3 that there are more traces of the marking medium in the wipe-off area suggesting that the contact of the tongue in the area of contact was not quite firm. The non-wiped off area in zones 3 and 4, shown in these four palatograms, exemplifies clearly the phonetic nature of sibilant fricatives. It indicates the region where the tip or blade of the tongue does not come in firm contact with the palate so as to form a narrow passage along the tongue midline through which the air escapes, thus producing the turbulance noise. Phonetically, for producing sibilant consonants, such as [s] and [ss], a groove is often formed along the tongue midline to channel the airstream. This is accomplished by touching the sides of the tongue to the side teeth.[1]

Palatograms 5-8 illustrate the articulations of single and geminate plosive consonants. They are of the words /ʕada/, /ʕadda/, /hadaf/ and /haddaf/, respectively. In these palatograms we find that the marking medium is clearly removed from the same zones as those of palatograms 1-4, i.e. the alveolar and the post-alveolar regions as well as all zones at the sides. But, whereas the wipe-off is mainly observable in the alveolar and post-alveolar regions for the fricatives, it greatly concentrates in the alveolar region for the plosives. There is a noticeable difference between the amount of wipe-off in palatograms 6

---

[1] Cf. Borden and Harris, 1980, p.121.

and 8 and that in palatograms 5 and 7. It is clearer and sharper in the former two than that in the latter ones. This, again, suggests that the tongue tip or blade touches the alveolar and post-alveolar regions more firmly when producing the geminate plosive [dd] than when producing its single counterpart [d] in word-medial position. Unlike palatograms 1-4, the non-wiped off area existing in zones 3 and 4 is missing in palatograms 5-8. This is due to the phonetic nature characterizing the production of alveolar and denti-alveolar plosive consonants in general. The wipe-off observable in zones 3 and 4 of palatograms 5-8 is brought about by a complete and firm contact between the tongue tip or blade and the alveolar ridge. These palatograms also show that the amount of wipe-off around the rim of the teeth, especially between zones 4-6, is apparently greater with the geminates than that with the non-geminates. This fact suggests that [dd] is accompanied by a firmer tongue-palate contact than [d], not only in the alveolar and/or denti-alveolar zones but also in some other palatal zones, i.e. in the pre-palatal and palatal regions.

5.4.2 Palatograms illustrating articulations of word-initial consonants

The same phonetic contrasts examined in palatograms 1-8 are now studied in word-initial positions. Palatograms 9-12 are of the words /saaξi/, /ssaaξi/, /suħub/ and /ssuħub/, respectively. In these palatograms, there is, again, a considerable difference between the removal of the marking medium observable in palatograms 10 and 12, and that in palatograms 9 and 11. The wipe-off is clearer and sharper in palatograms 10 and 12 than that in the other two palatograms indicating a firmer contact of the articulatory organs, namely the tongue and the palate, during the articulation of the geminate [ss] than during the articulation of the non-geminate [s] pronounced word-initially. The non-wiped off area seen in zones 3 and 4 of palatograms 1-4, which characterizes the distinctive nature of the production of alveolar fricatives, can also be observed in these palatograms. Furthermore, there is a clearer and more extensive removal of the marking medium from zones 5-7 during the articulation of the geminates than during the articulation of the non-geminates. Palatograms 9 and 10 show a considerable wipe-off between zones 6 and 7. These are probably the zones where the articulating organs come into contact during the enunciation of vowel [i] occurring word-finally in /saaξi/ and /ssaaξi/.

Palatograms 13-16 illustrate the articulation of single and geminate plosive consonants in the words /daaξi/, /ddaaξi/, /digam/ and /ddigam/, respectively. These palatograms show the least difference of wipe-off resulting from the articulation of geminate consonants and that resulting from the articulation of their single partners. Nevertheless, a slightly clearer and more extensive removal of the marking medium is noticeable in zones 3 and 4 of palatograms 14 and 16, suggesting a firmer tongue-palate contact during the articulation of the geminate [dd] than during the articulation of its single cognate [d] pronounced word-initially.

Similar to that of palatograms 9 and 10, the considerably greater removal of the marking medium observable in zones 6 and 7 of palatograms 13 and 14 is almost certainly caused by the production of vowel [i] occurring word-finally in /daaξi/ and /ddaaξi/. In palatograms 15 and 16, on the other hand, it probably takes place as a consequence of producing [g] used inter-vocalically both in /digam/ and in /ddigam/.

5.4.3 General

The idea of the zones, as described by Firth (op.cit.) was suggested primarily to introduce precision and quantitativeness to palatographic measurements for the sake of inferring the degree of contact between the articulating organs. Therefore, an attempt was made to measure the areas from which the marking medium had been removed. This was accomplished by applying a transparent overlay on each of the sixteen palatograms. The transparent overlay was a xeroxed copy of a graph paper with small squares. Each square was 2.5 mm in size.

Since our present experiment was mainly concerned with investigating the degree of contact between tongue and palate while producing single and geminate consonants formed at the alveolar and/or denti-alveolar regions, our actual measurements were made only between zones 2-4, shown on each palatogram. The number of squares was counted at places where the marking medium was missing. The measurements obtained are shown in table 5.2 below.

Table (5.2)

Tokens	words with single consonants	Number of squares	words with geminate consonants	number of squares
1	/ξasa/	48	/ξassa/	56
2	/ħasab/	55	/ħassab/	64
3	/ξada/	47	/ξadda/	79
4	/hadaf/	73	/haddaf/	86
5	/saaξi/	54	/ssaaξi/	61
6	/suħub/	52	/ssuħub/	65
7	/daaξi/	77	/ddaaξi/	82
8	/digam/	79	/ddigam/	81
	Mean	61	Mean	72
	S.D.	13.4	S.D.	11.4
	T		3.413	
	p		0.0112*	

The above measurements show very clearly that words with geminate consonants have more squares of wipe-off than those with single consonants. This denotes that geminate consonants are characterized by having a more extensive wipe-off than their single partners, thus suggesting firmer tongue-palate contact during the enunciation of the geminates than that during the enunciation of the non-geminates. These measurements were subjected to the statistical treatment of a two-tailed t-test for related samples, i.e. for matched-subjects design,[1] so as to find whether the number of squares

---

[1] For further details about the t-test, see section 6.11.

associated with geminate consonants is significanly different from that associated with single consonants. The results shown in table 5.2 indicate quite obviously that the difference is statistically significant, where p value is at 0.0112.

5.5 Experiment No. 2 (Electropalatography)

Experiment No. 1 was updated by employing electropalatography (EPG) in this experiment. Since the EPG has been quite recently introduced into the Phonetics Laboratory, the computer program used for data processing and for producing graphic printouts is still in preparation. It will probably take some more time to be in its final form. Once this program is completed, the EPG will then be ready for full exploitation in sophisticated phonetic research.

The present investigation was only a pilot experiment which aimed at confirming the direct palatography findings. It also aimed at discovering the manoeuvres of the articulating organs during the production of single/geminate consonants in I.C. Arabic.

5.5.1 Experimental Procedure

5.5.1.1 Selection of Stimuli

Only two minimal pairs were used in this experiment. They were selected from those studied in Experiment No. 1. The two pairs are: /ξasa/ versus /ξassa/, and /ξada/ versus /ξadda/. The examined segments represent two oppositions of different contoid types where contrasts between word-medial single fricative or plosive and their geminate cognates were involved. (See section 5.2.1.1 for details). All test items were uttered in isolation at normal speed.

5.5.1.2 Subjects

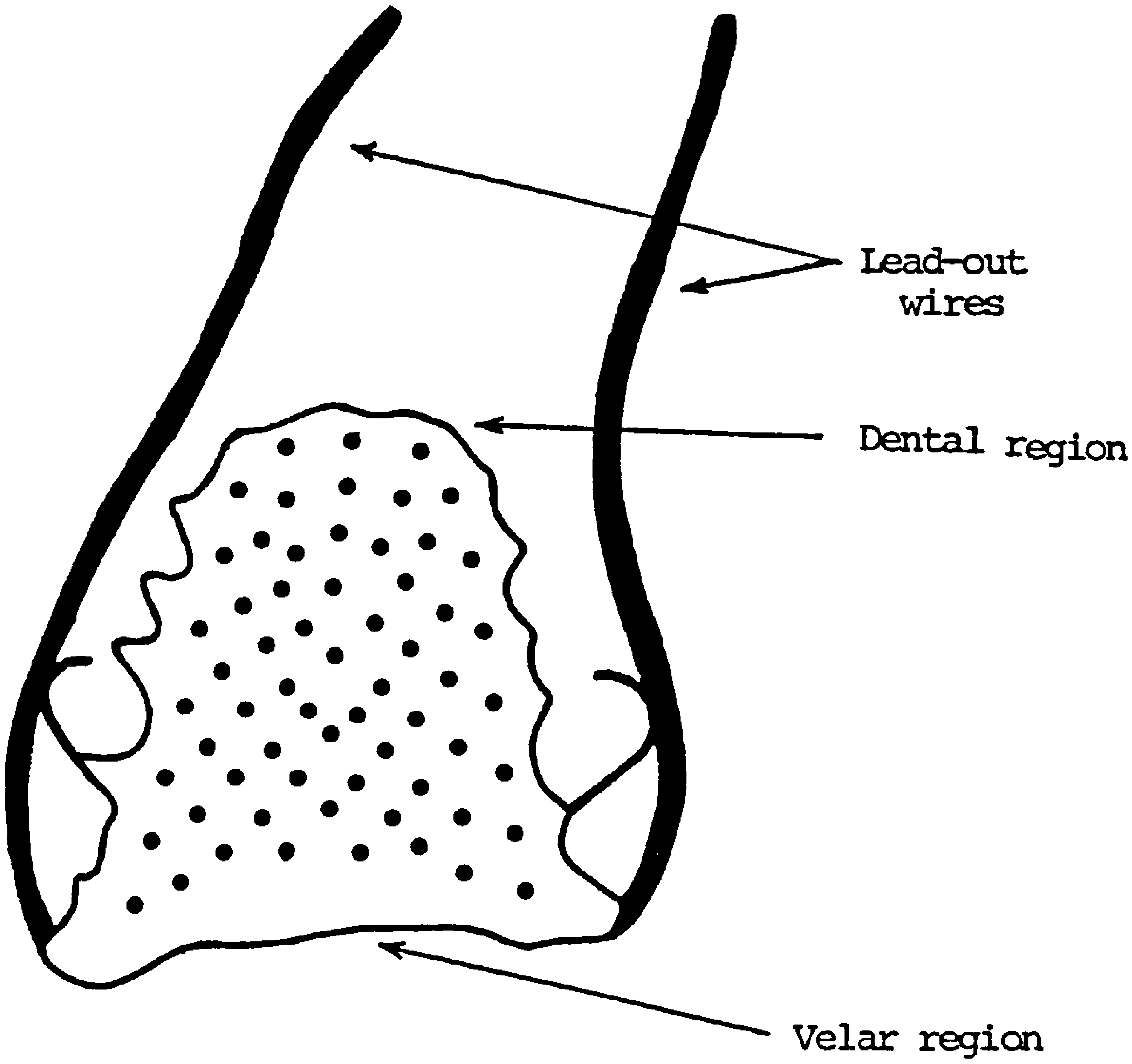
The experimenter was his own subject in this experiment also.

5.5.1.3 Instrumental Set-up and Method

A 63-contact artificial palate was made to fit the roof of the subject's mouth. This palate can be connected to the EPG via two separate lead-out wires. The EPG (model DP-01) is a relatively small portable device manufactured by Rion Company Ltd., Tokyo, Japan. It is operatable by means of four D-size batteries. A small screen and the functional switches and buttons are all located on the front side of the device. The screen contains 63 small red bulbs representing 63 electrodes embedded in the artificial palate. The arrangement of electrodes on the subject's artificial palate is shown in Fig. 5.3. An additional, and relatively larger, electrode is adjusted on the outside domed surface of the artificial palate to make a neutral contact with the roof of the mouth. Below the screen are the 'on-off' switch as well as the 'memory', the 'stop' and the 'single step' buttons.

The tongue-palate contact can be clearly seen on the screen once the subject inserts the artificial palate into

Fig. (5.3) The 63-contact artificial palate used in the EPG experiment shown here fitted to a plaster impression of the subject's palate and teeth.



his mouth, provided that the machine is switched on. Whenever the tongue touches the palate, some of the small red bulbs will light up indicating contact between the tongue and the palate. All bulbs will light up when there is total contact between the two articulating organs.

The results of this experiment were obtained by first uttering the test items separately. Once the test item was produced, a visual display was shown on the small screen indicating contact. By pressing the 'memory' button immediately after uttering the test item, the original display would be stored in the memory of the machine. The red bulbs would continue displaying the tongue-palate contacts of the uttered test item until the 'stop' button was pressed. This would freeze the continuous lighting up of the red bulbs to indicate one single display (or frame). For careful inspection, the subject could retrieve the overall tongue-palate contacts made during the production of the test item simply by pressing the 'single step' button. Each time the subject pressed this button he would get one single frame at a time on the screen. When the frames came to an end, a small red bulb at the top right-hand corner of the screen would light up.

The subject had a selection of three different buttons from which he could choose the appropriate time he would like for examining successive displays. He could either press the first button for normal time display, or he could press the second or third buttons for three times slower or ten times slower than normal time display, respectively.

Each successive display represented an interval of sixty fourth of the second, i.e. 15.6 msec. In this experiment, the normal time for displaying the frames was used. All graphic displays representing lingual-palatal contacts were drawn by the experimenter. Solid circles indicate contact; empty ones indicate no contact.

5.6 Analysis of Results

Figs. 5.4-5.7 show various EPG displays illustrating variations in the tongue-palate contact between a single consonant and a geminate counterpart. It is evident from these displays that the geminate consonants are characterized by a considerably greater number of area contacts than the non-geminate consonants. Figs. 5.4 and 5.6 show fairly typical patterns for [s] and [d]; whereas Figs. 5.5 and 5.7 show the expected patterns for [ss] and [dd] respectively.

Further examination of these figures also shows a different build-up in the contacts when the singles are compared with the geminates. For instance, the maximum number of contacts achieved during the occlusion phase of a geminate plosive is considerably greater than that achieved during the occlusion phase of a single partner. Fig. 5.7 shows very clearly that there is total contact between the articulating organs during the production of a geminate plosive. Whereas the maximal area contact for a geminate consonant lasts for a noticeably greater number of frames, it only lasts for a relatively smaller number of frames for a single cognate. This strongly suggests that a geminate consonant is significantly longer than a non-geminate one, as is clearly shown in Figs. 5.8 and 5.9.

The results of this pilot experiment are highly consistent with those of the direct palatography in that a geminate consonant is accompanied by firmer lingual-palatal contact than a single counterpart. They are also consistent with our previous experiments in that geminate consonants are considerably longer than their single partners. The very extensive contact indicated by the EPG is unexpectedly different from the direct palatography pictures.[1]

5.7 Discussion

The findings of these experiments lend support to the view that during the articulations of fricative and plosive consonants pronounced either in word-initial or in word-medial positions, the tongue touches the area of contact more firmly during the articulation of the geminate consonants than during the articulation of their single counterparts. These results seem to be in complete agreement with previous findings reported by other investigators.

Thananjayarajasingham (1976), for example, found that the results of his palatographic study indicated that during the articulation of 'double' nasals and laterals in Ceylon Tamil, there was a firmer tongue-palate contact that during the articulation of single nasals and laterals. The firmer contact between the tongue and palate resulted in a clearer and more extenisve wipe-off for the 'doubles' than for the singles. The doubling, i.e. the lengthening, of intervocalic lateral and nasal consonants is said to be 'distinctive' in Tamil. Thananjayarajasingham's findings have been recently

---

[1] This could be due to a faulty threshold setting causing salival bridging.

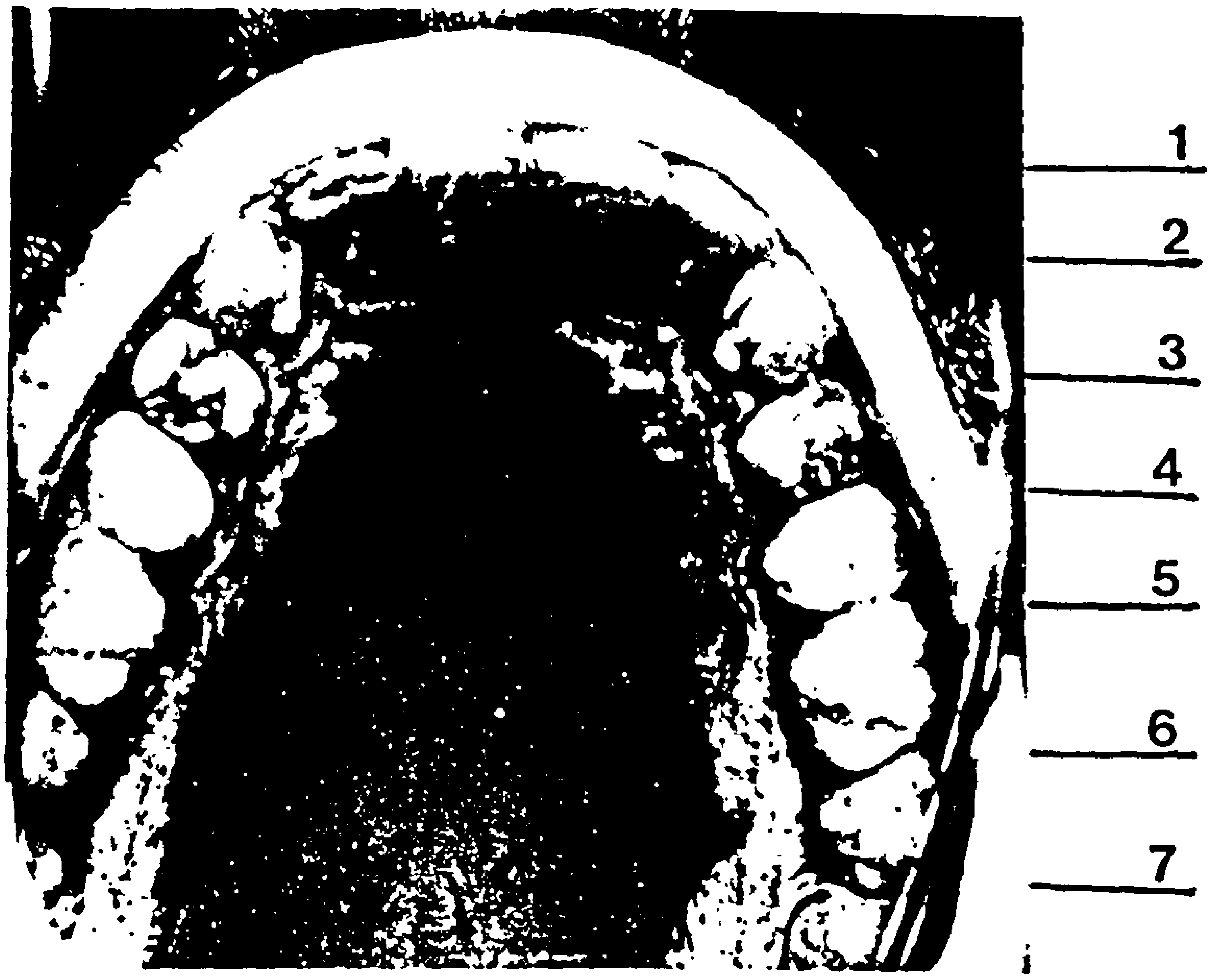
confirmed by Balasubramanian (1982) whose palatographic and kymographic investigation of the doubling of intervocalic nasal and lateral consonants in Indian Tamil showed that there was a clearer and more extensive removal of the marking medium when producing 'double' nasals and laterals than when producing their single cognates. This indicates that the articulating organs, viz. the tongue and the palate, make a firmer contact with each other during the articulation of the 'double' consonants than during the articulation of their single partners.

In addition, McGlone and Proffit (1972) conducted a dynamic palatographic investigation in which they employed "a system using strain gage pressure transducers as sensing elements." [p.375]. Using this system, the two researchers tried to show 'mechanical pressure', i.e. the force of how the tongue pushes against the roof of the mouth during speech production. Their results showed that greater pressure was found for syllables containing [t] than those containing [d], and that the former syllables produced longer tongue-palate contact than the latter ones.

Our present findings can be interpreted by following the same lines suggested by McGlone and Proffit (ibid.). Articulatorily speaking, our data showed that there was some evidence indicating that there was a longer contact (i.e. occlusion) for the geminates than for the non-geminates. It is quite likely that a longer duration of contact for the geminates [ss] and [dd] will be associated with greater mechanical pressure and larger area of contact. That is, the clearer and greater wipe-off associated with the geminate consonants, in comparison with that of their single counterparts, was possibly a natural outcome of a greater pressure exerted by the tongue against the palate during the production of the geminates than during the production of the non-geminates. At the same time, our results seem to be consistent with the longer occlusion of geminate consonants reported in our acoustic experiment of Chapter Three.

Earlier in this study, it was shown that geminates are, acoustically, considerably longer than their single cognates whether they occur in word-initial or word-medial positions. The perceptual salience of differences in duration between single and geminate consonants was also demonstrated in our experiments using synthetic speech. The results of these experiments give further support to our previous findings. A geminate consonant has been found to be produced with a noticeably clearer and more extensive wipe-off which indicates a firmer contact between the articulating organs; whereas the single consonants are characterized by a less clear wipe-off. Accordingly, one may conclude that the clearer and the more extensive the wipe-off is, the firmer the contact between the articulating organs.

Palatogram
1



/ξasa/

Palatogram
2

/ξassa/

Palatogram .

3

/ħasab/

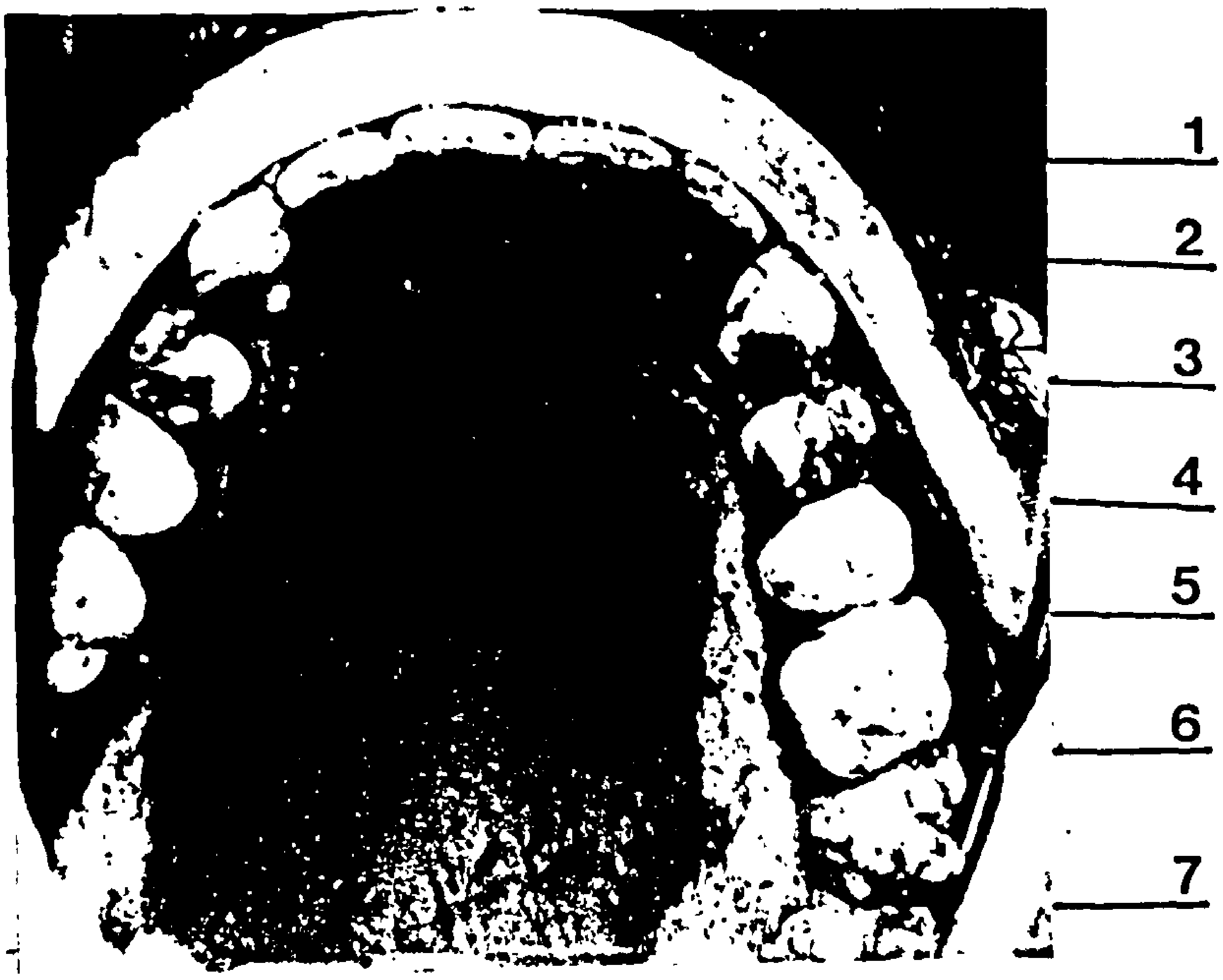
Palatogram

4

/ħassab/

Palatogram

5



/ʕada/

Palatogram

6

/ʕadda/

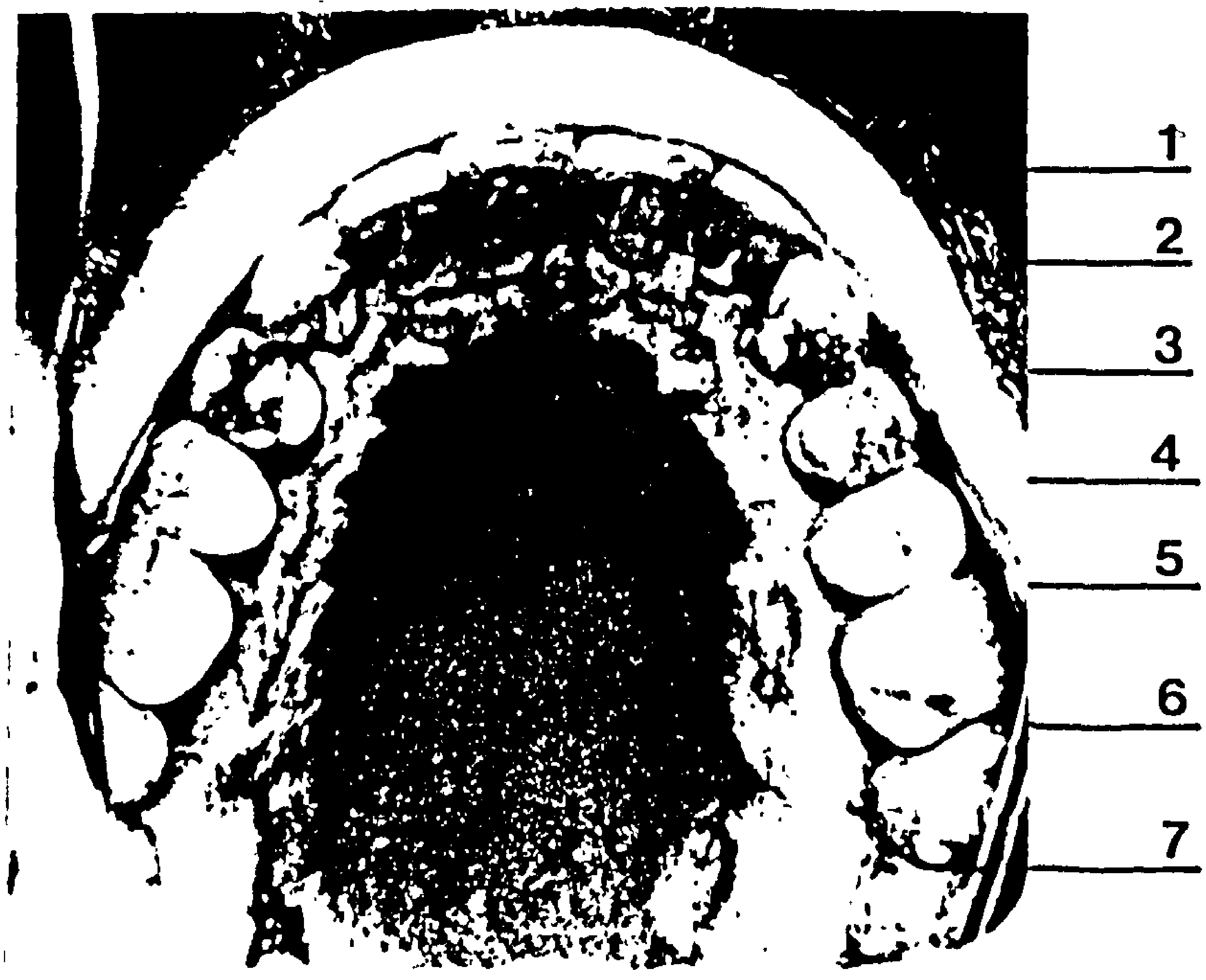
Palatogram

7

/hadaf/

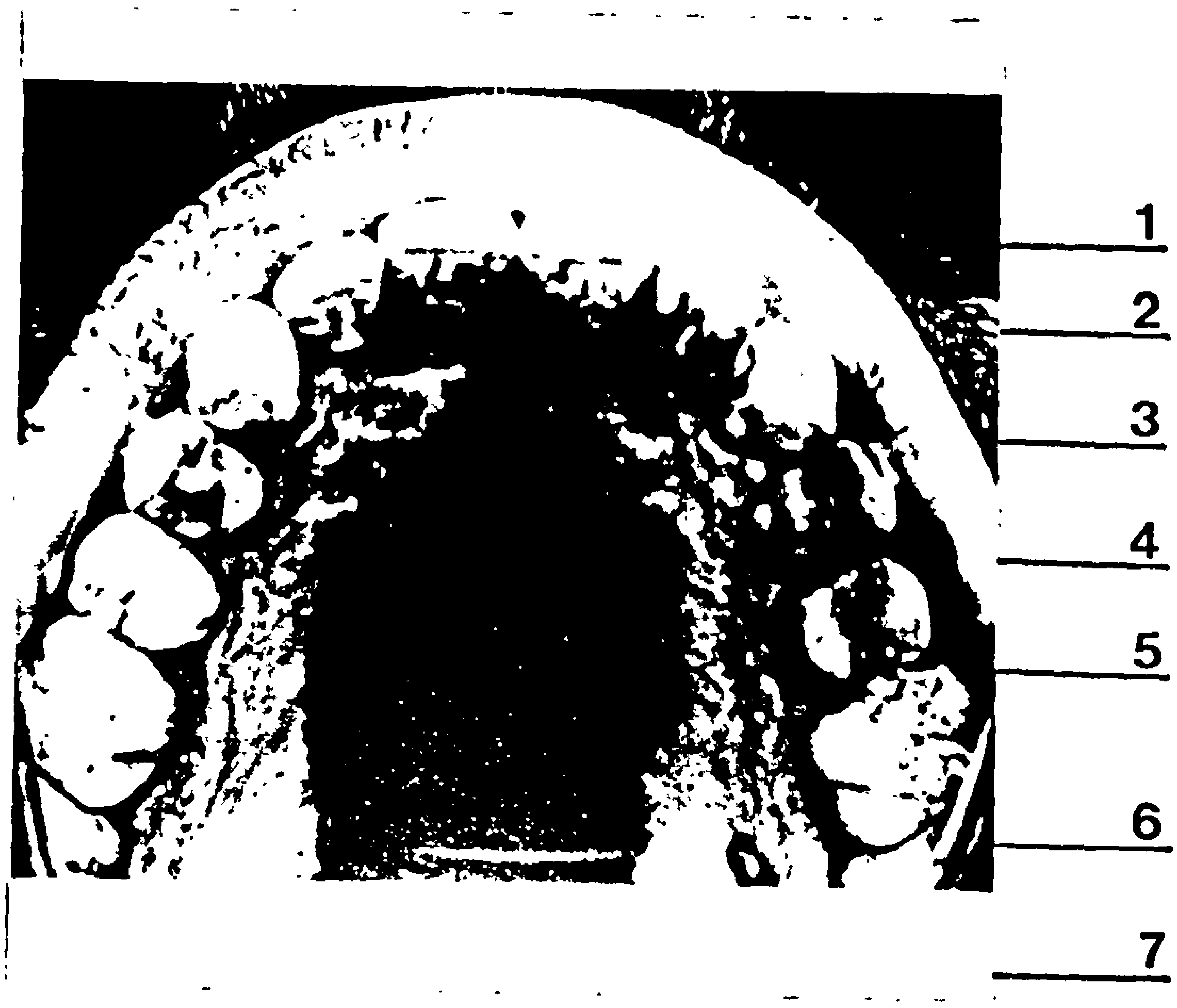
Palatogram

8



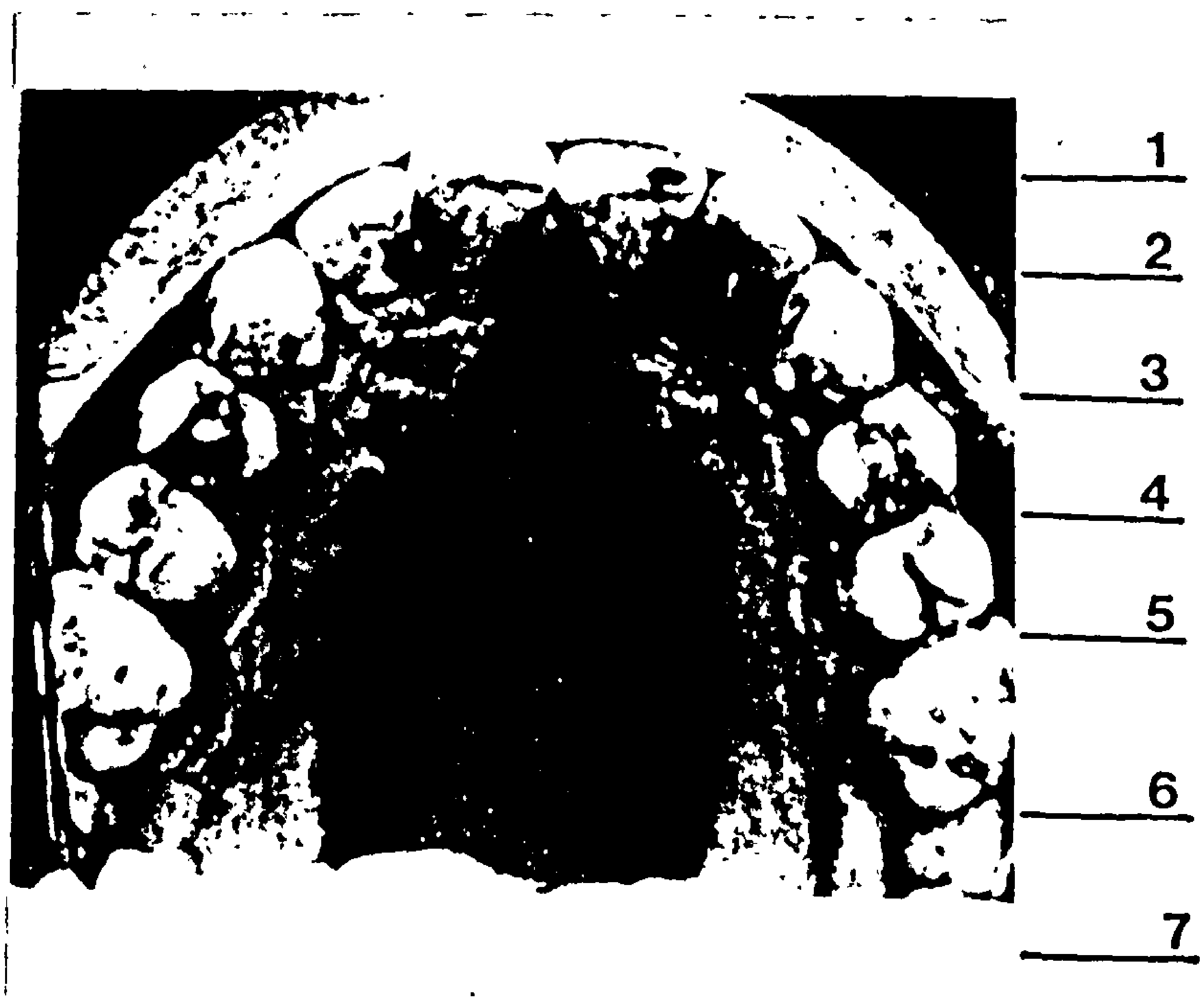
/haddaf/

Palatogram.
9



/saaξi/

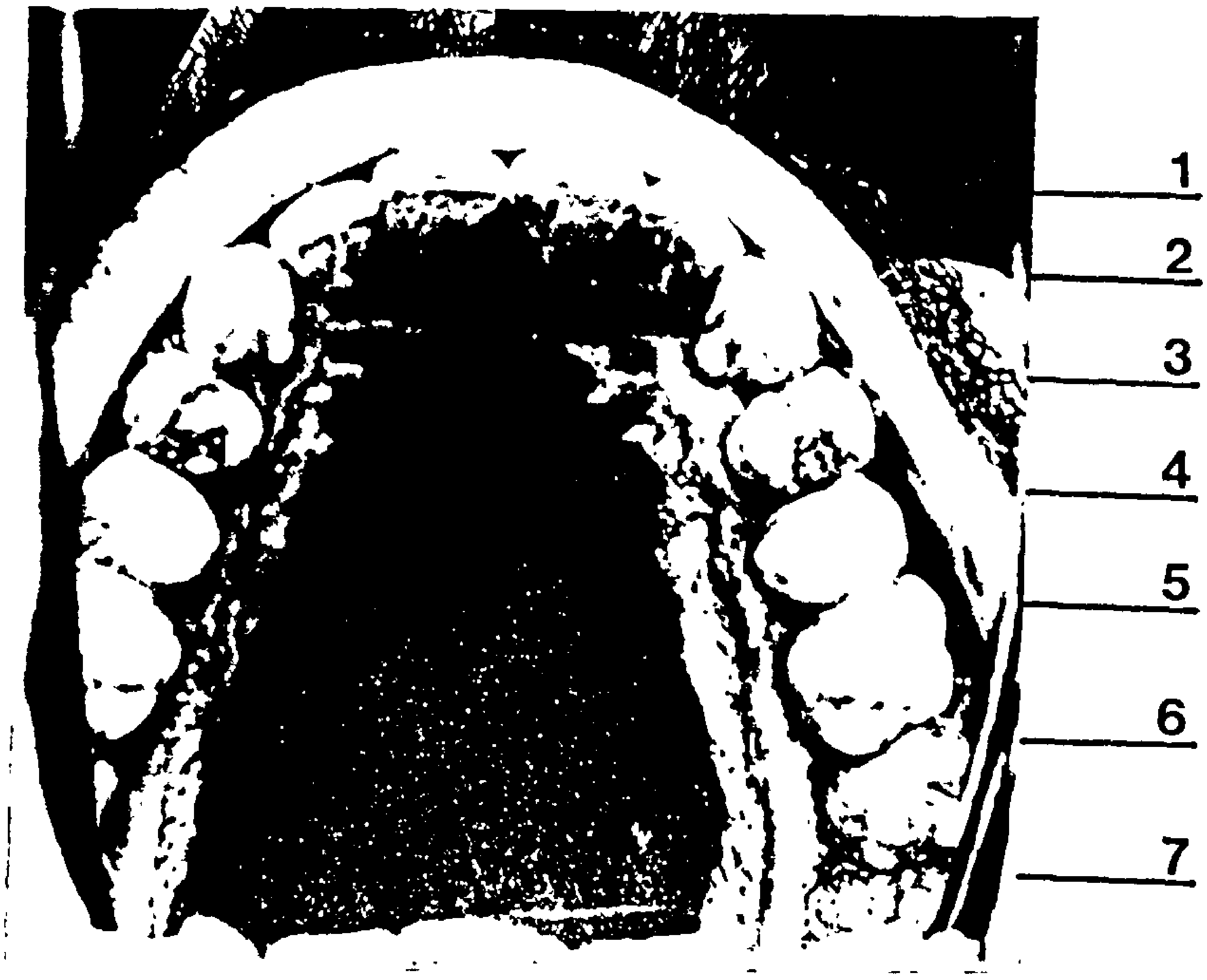
Palatogram
10



/ssaaξi/

Palatogram

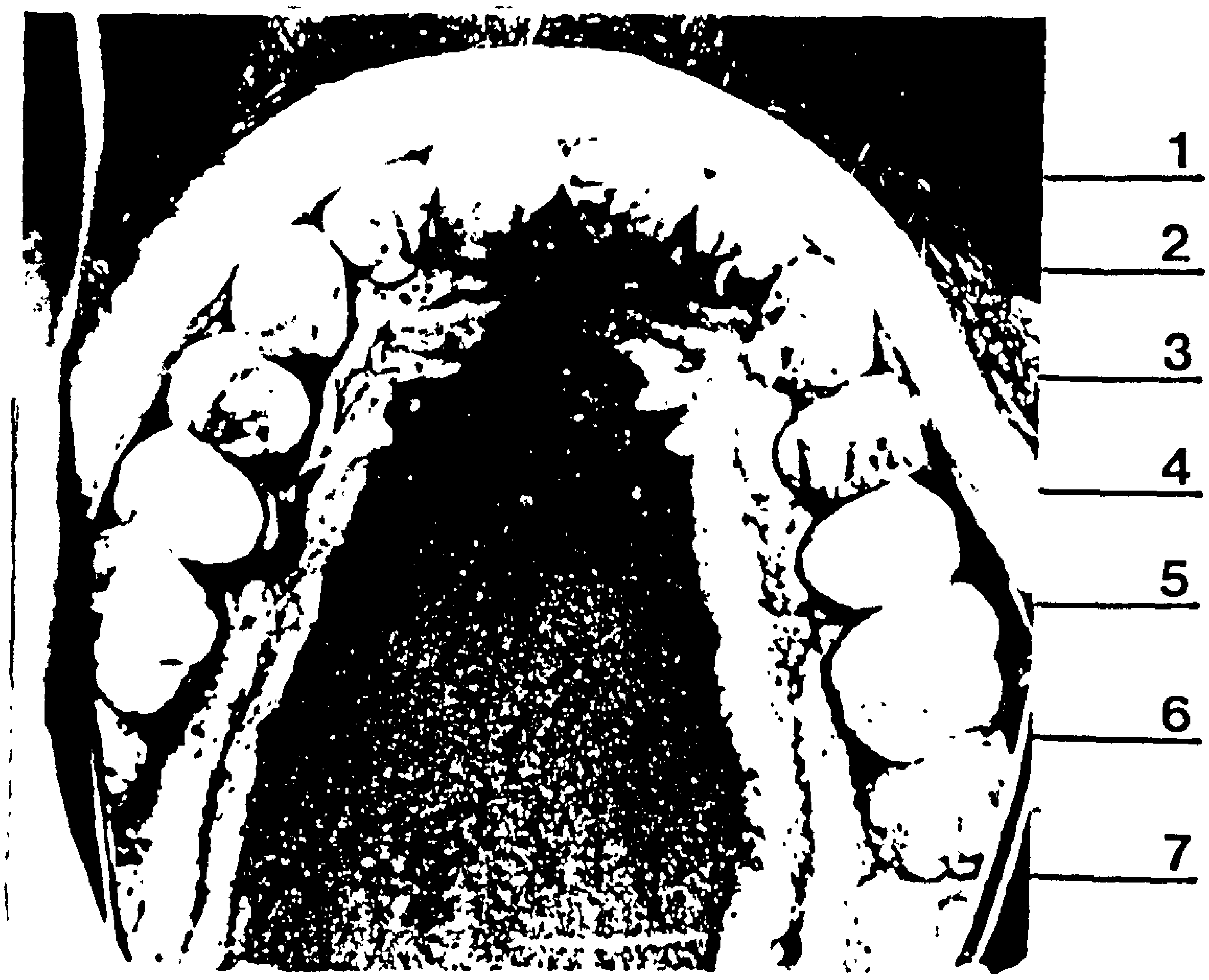
11



/suħub/

Palatogram

12



/ssuħub/

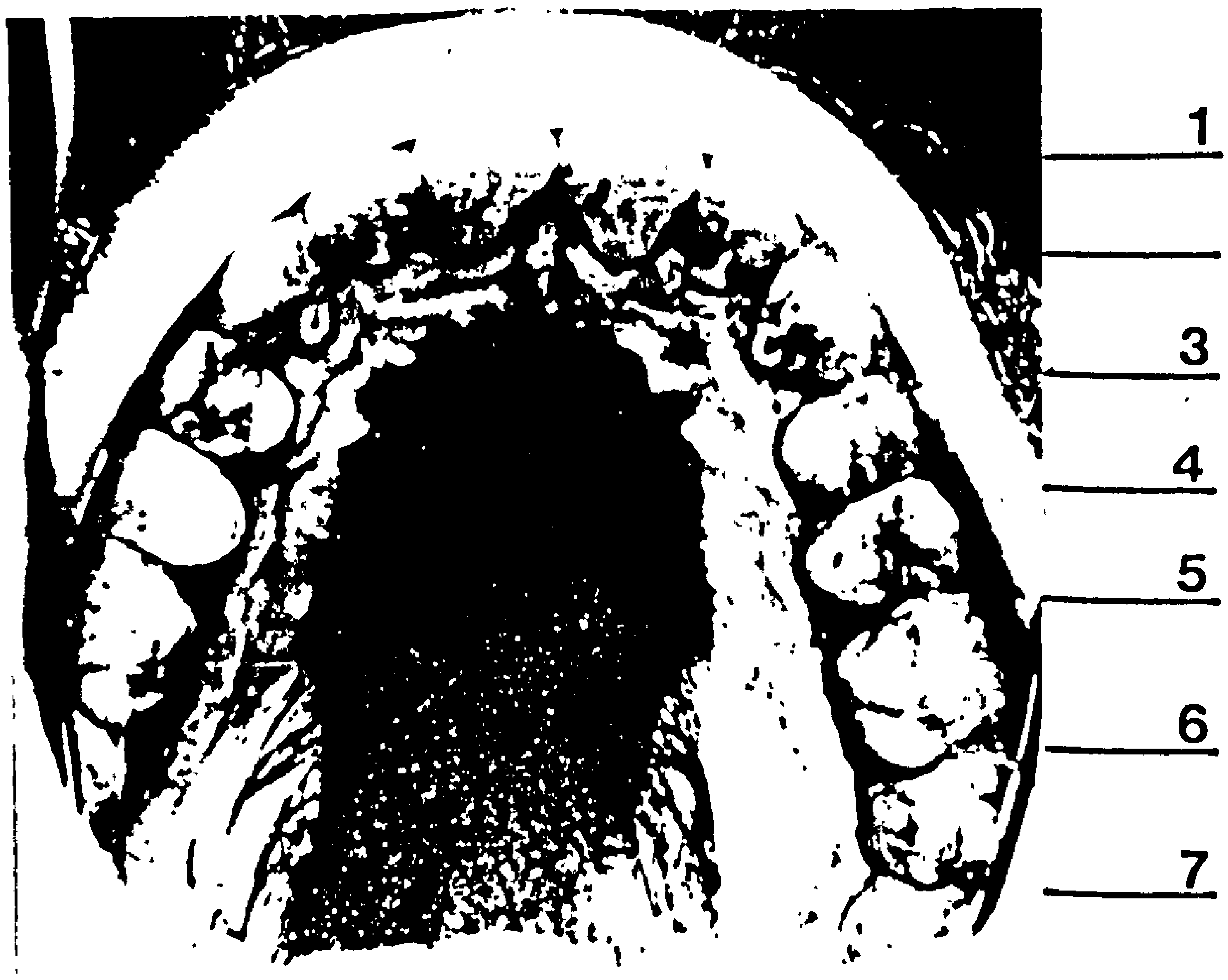
Palatogram

13

/daaξi/

Palatogram

14



/ddaaξi/

Palatogram

15

/digam/

Palatogram

16

/ddigam/

Fig. (5.4) Diagrams showing EPG patterns for the single fricative [s] in the environment of /ξasa/.

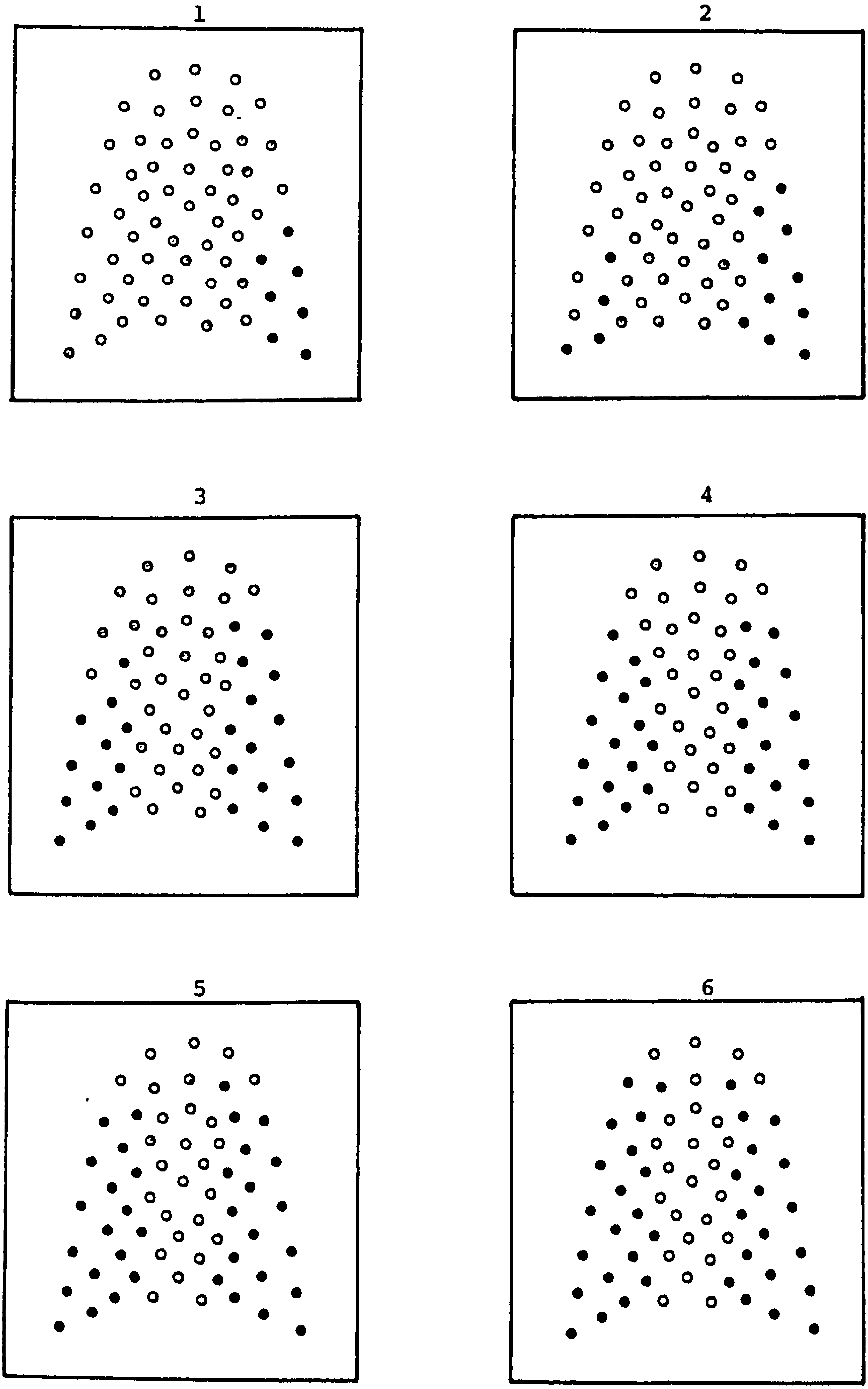
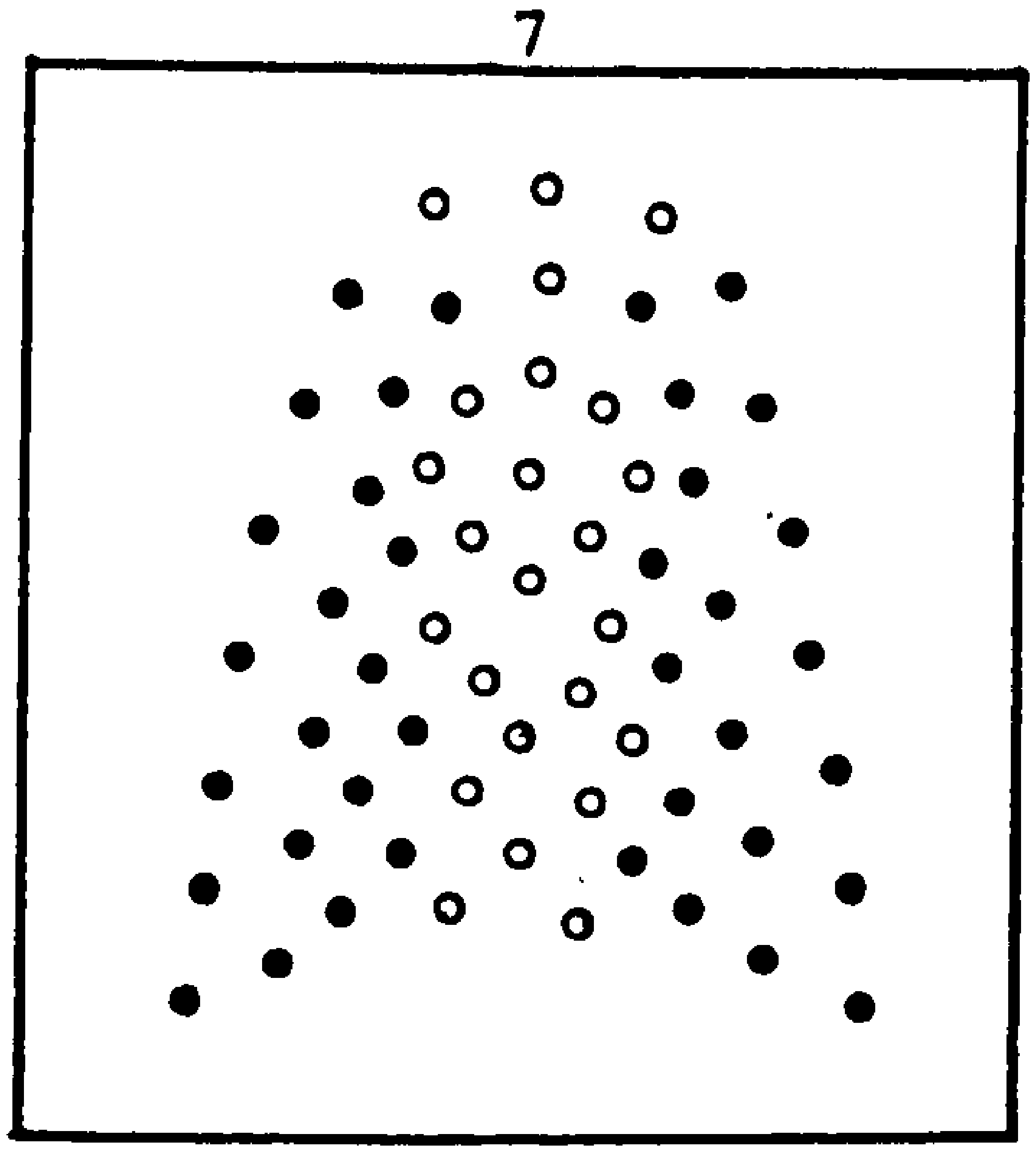
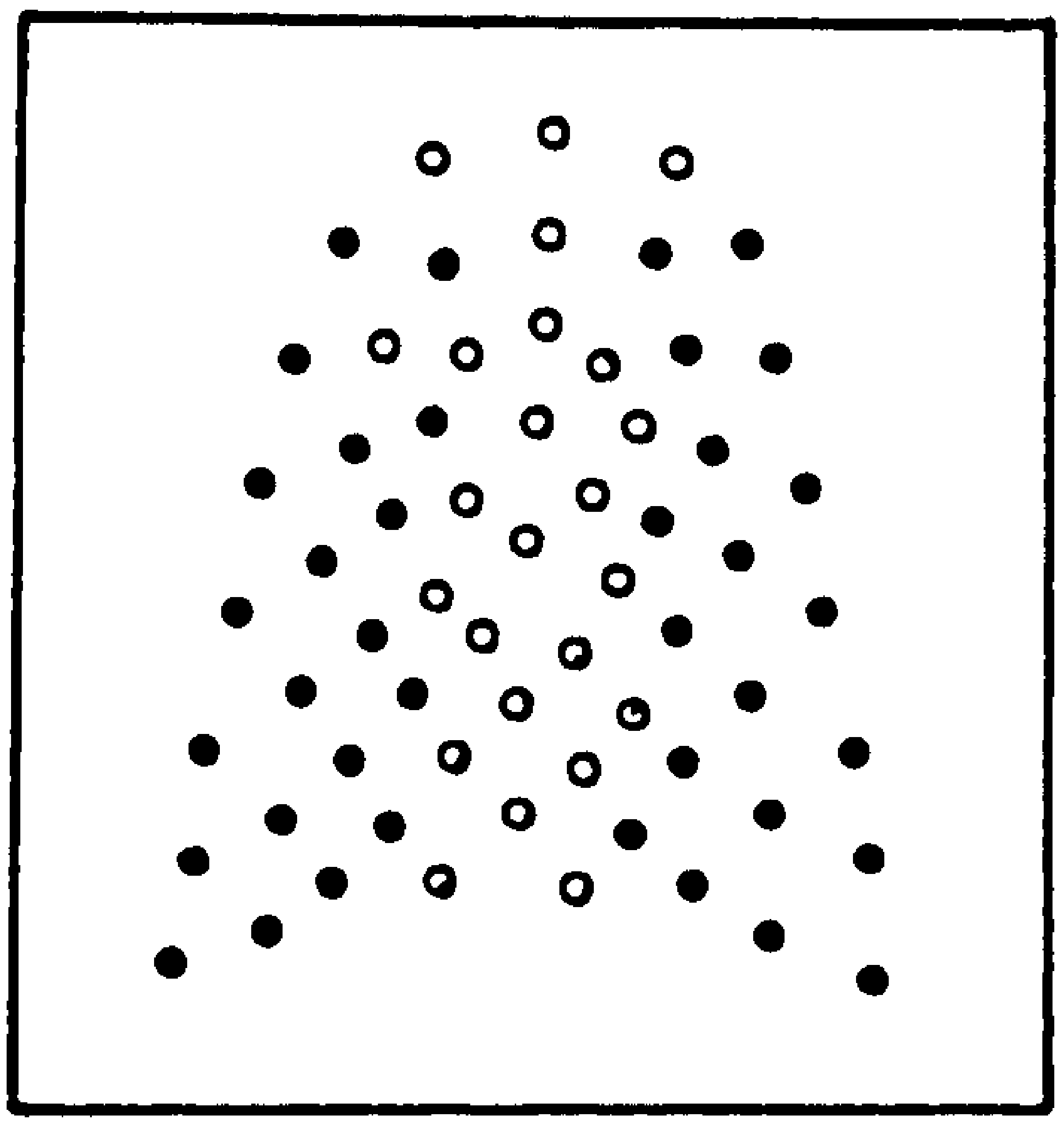


Fig. (5.4) cont'd

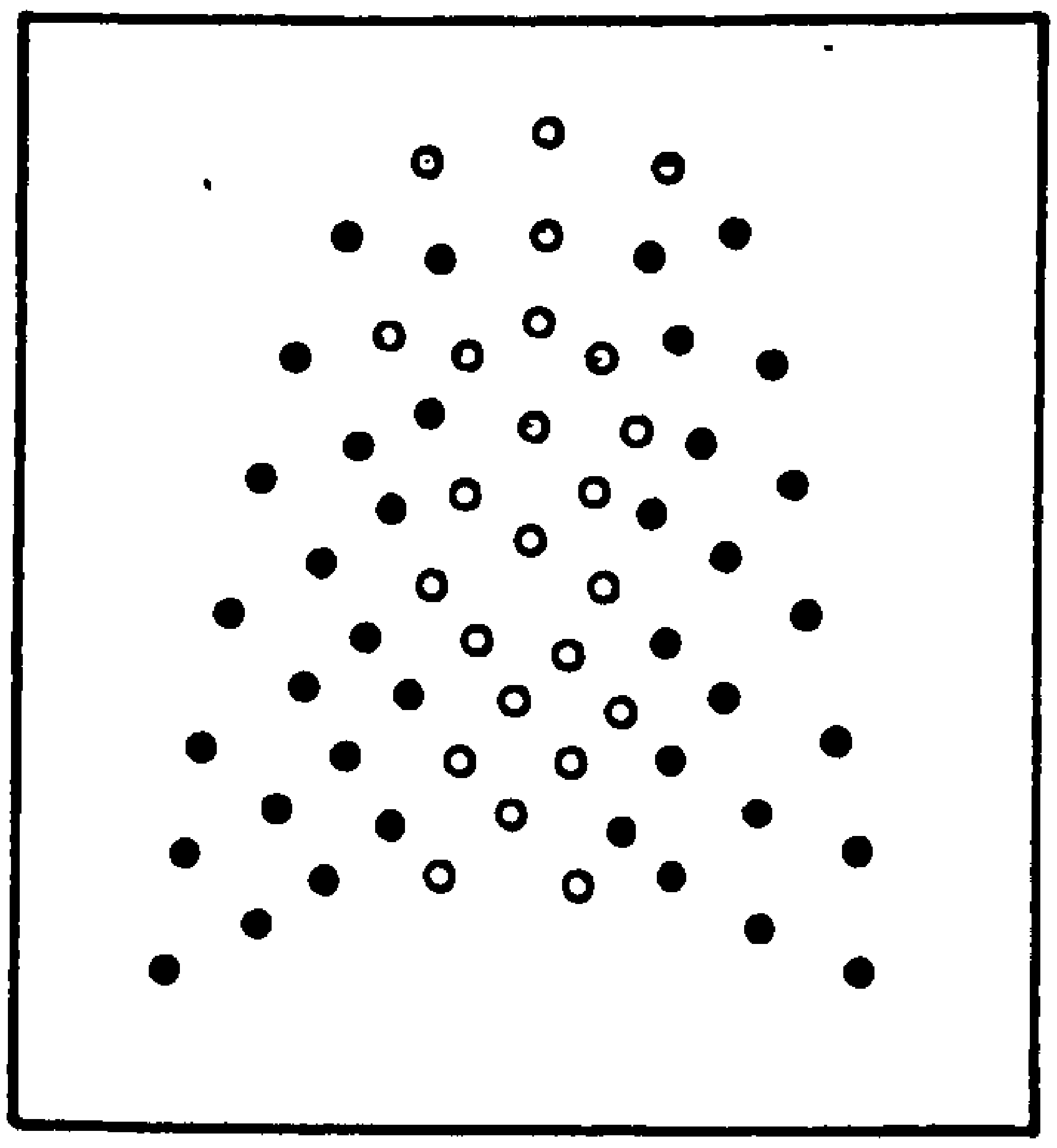
7



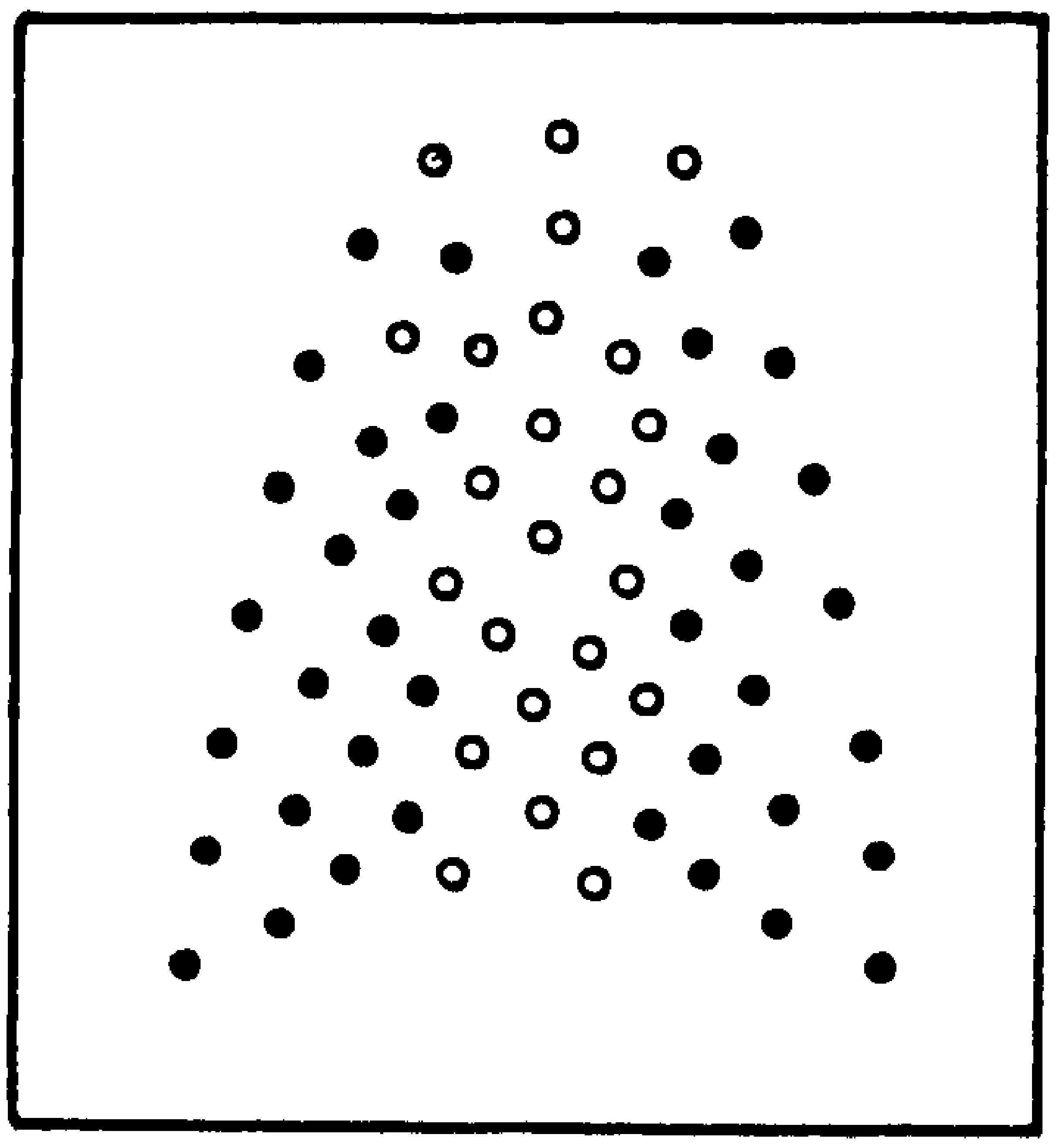
8



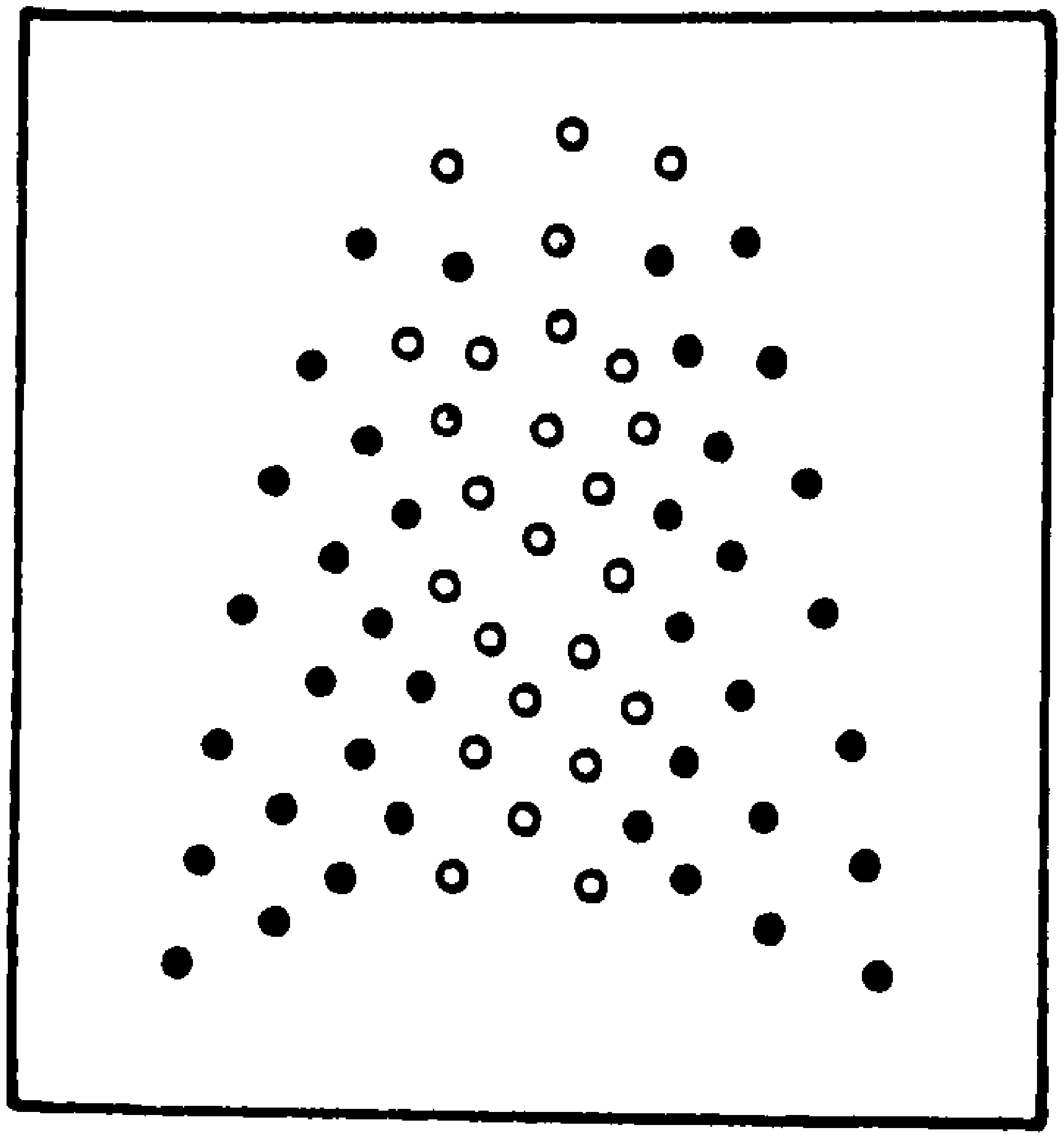
9



10



11



12

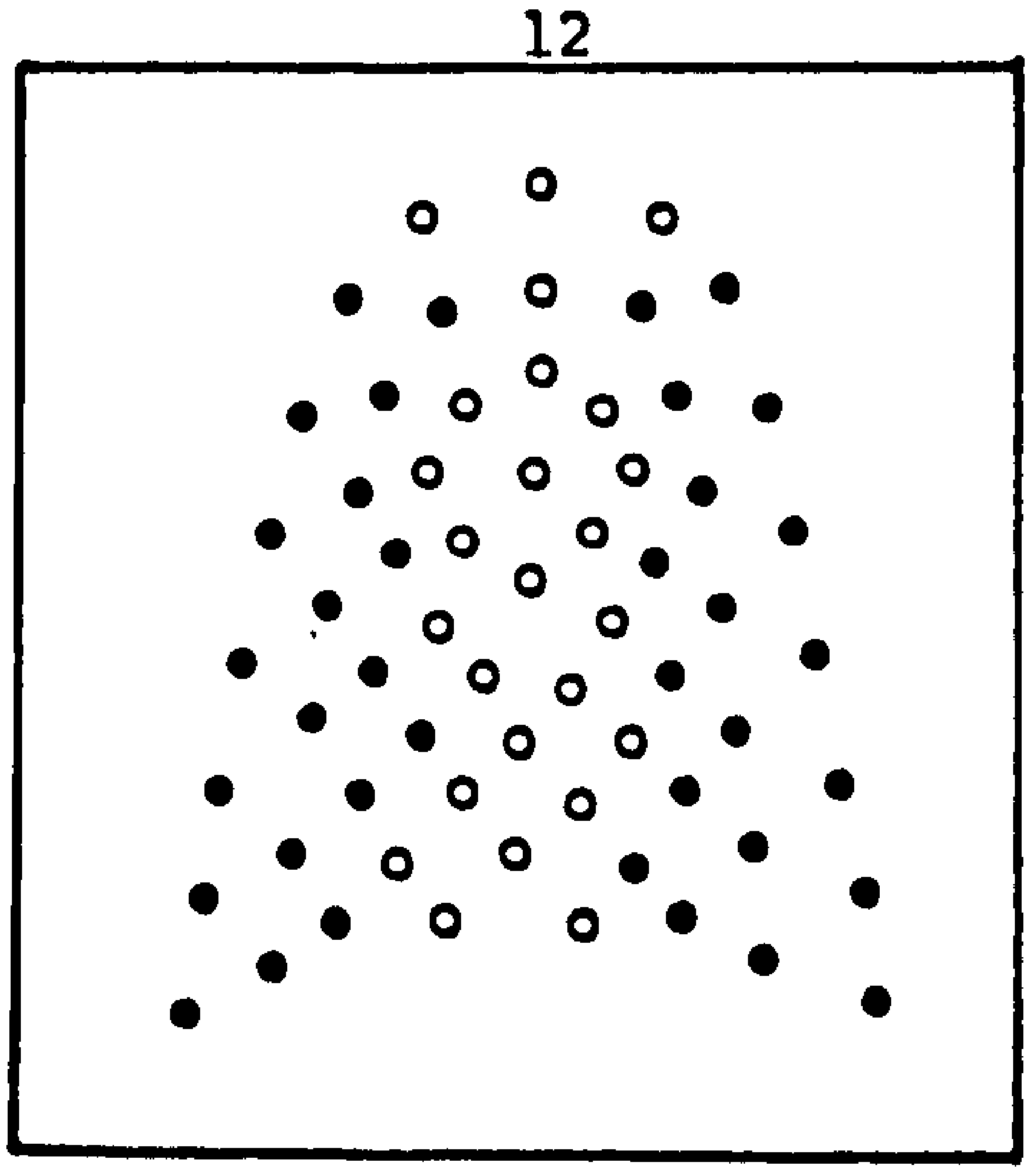
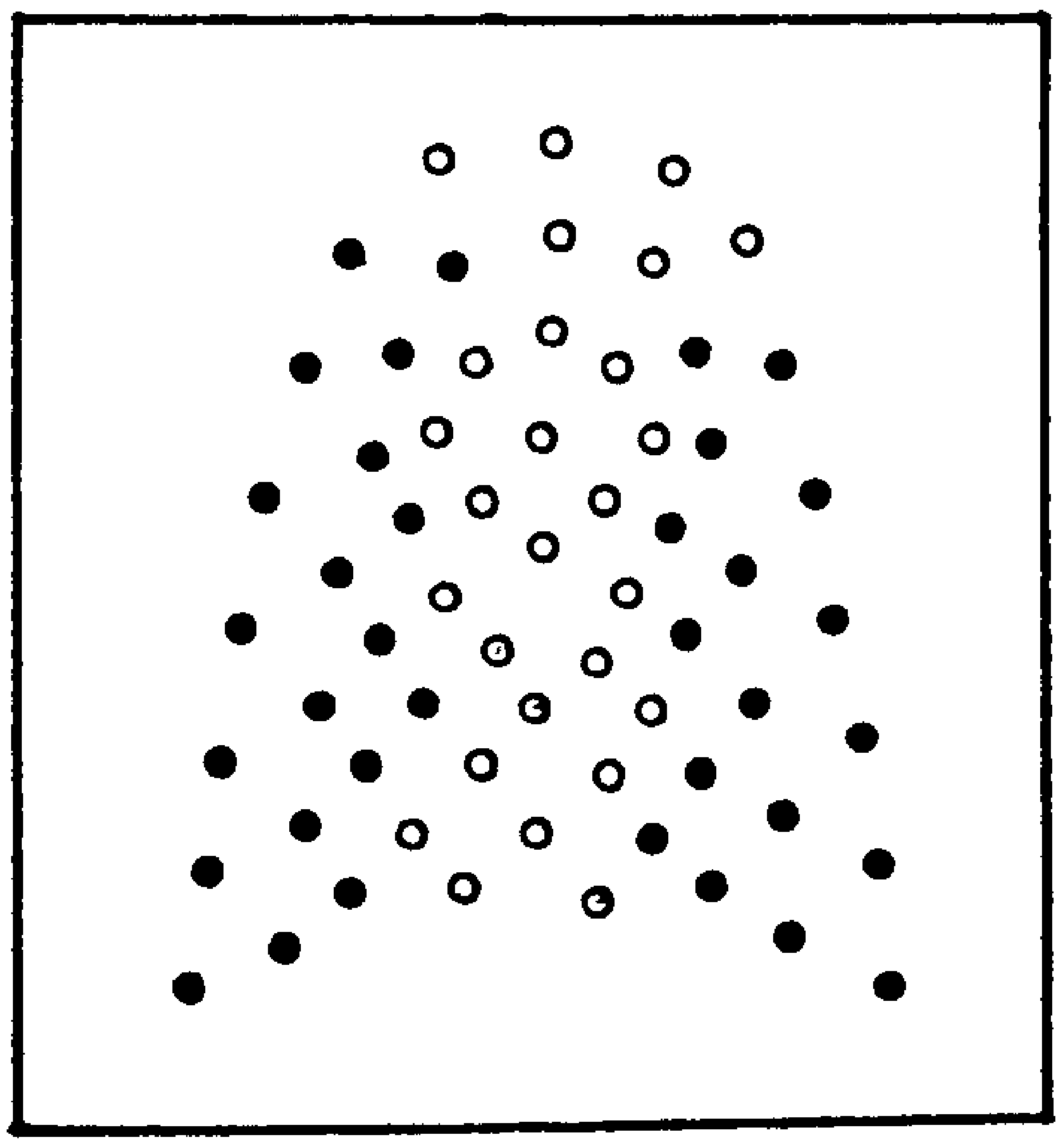
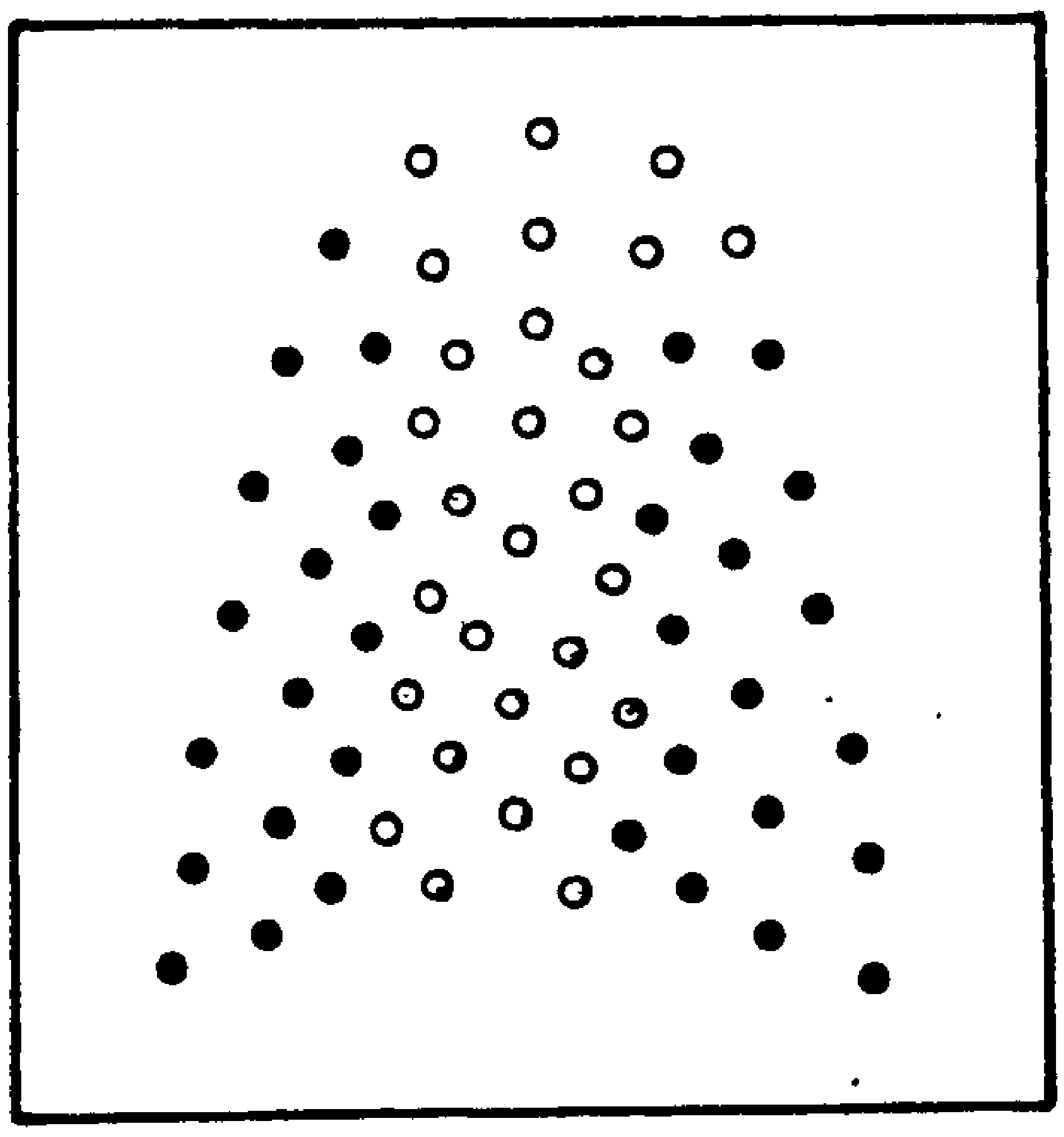


Fig. (5.4) cont'd

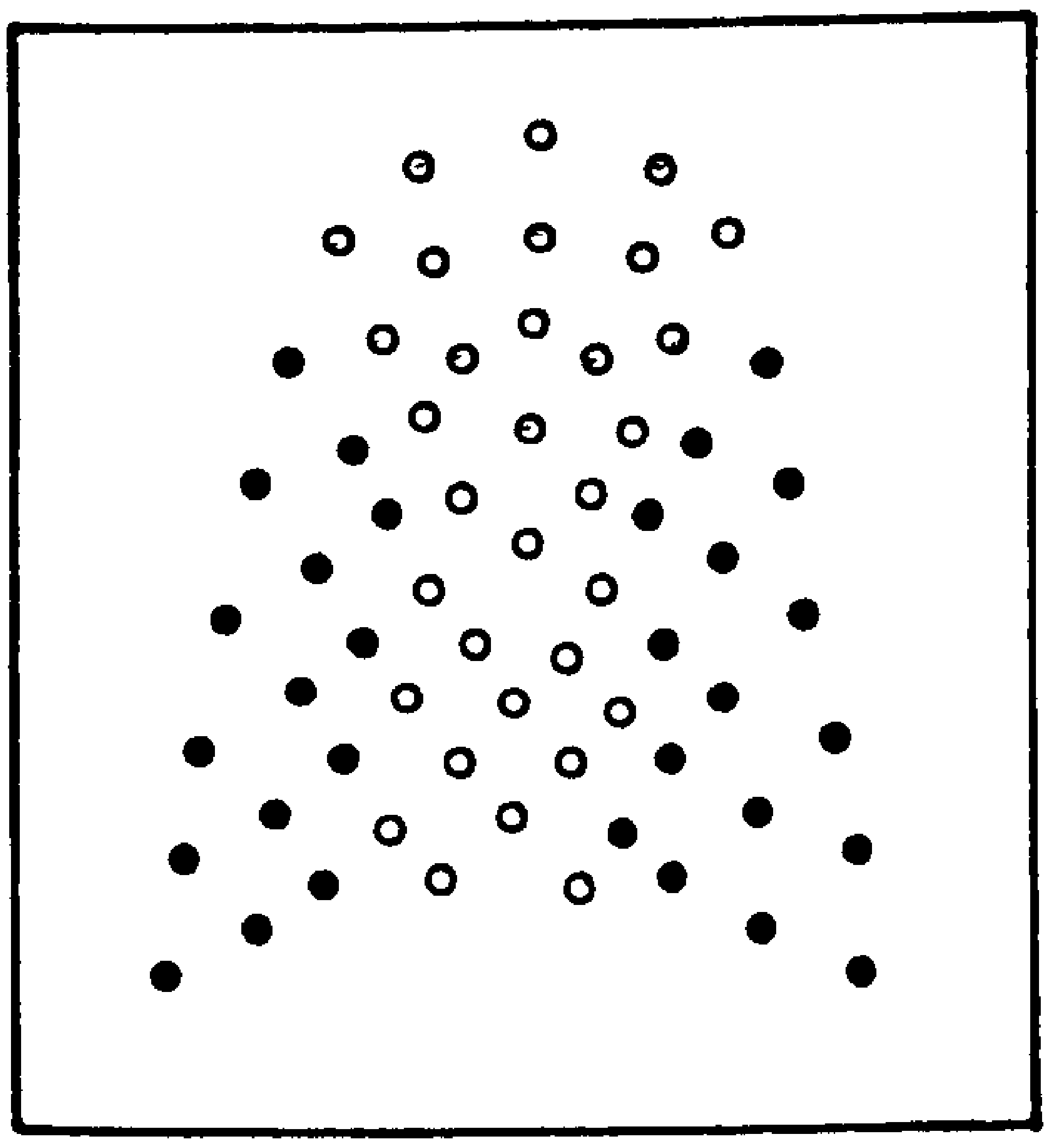
13



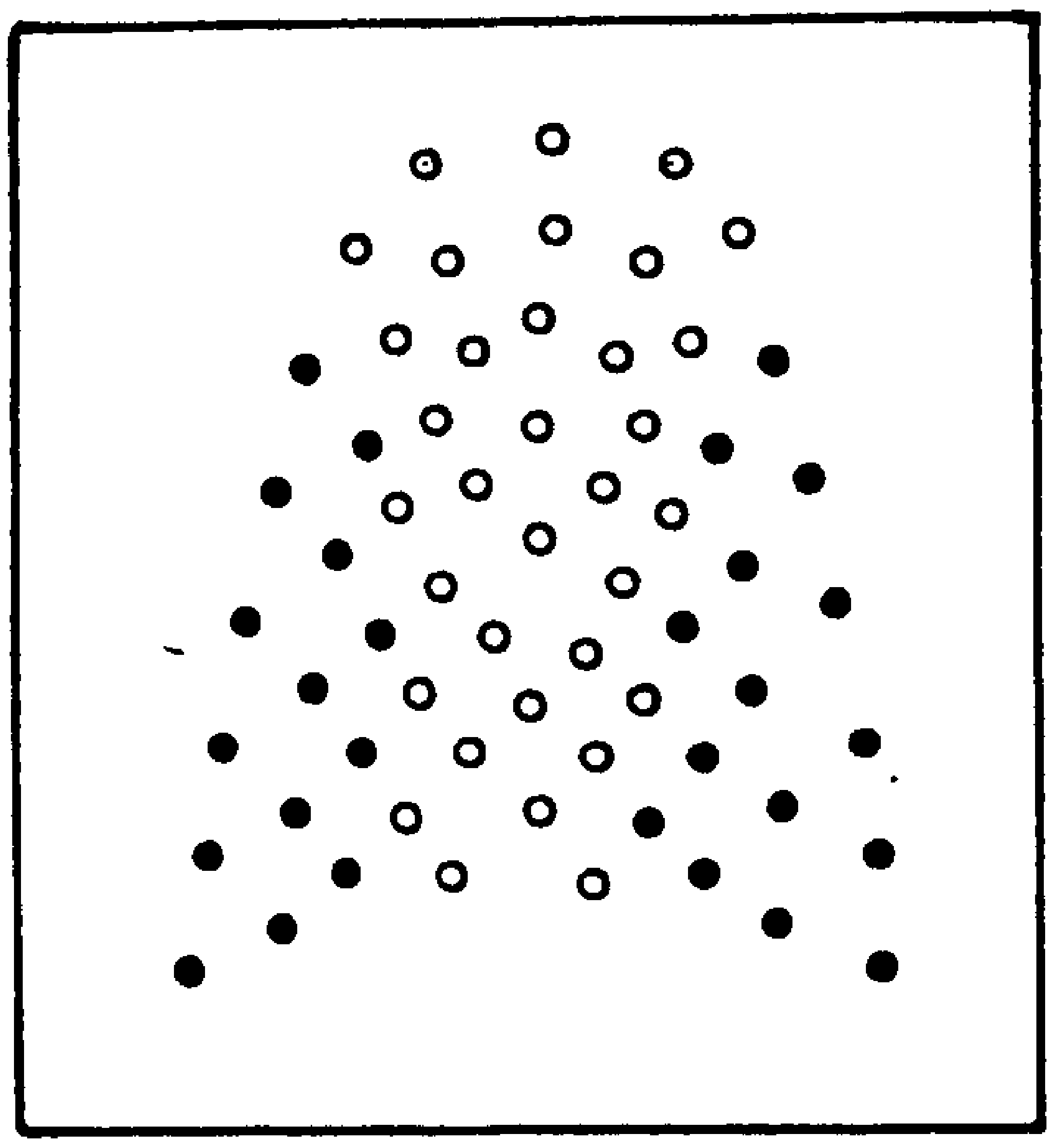
14



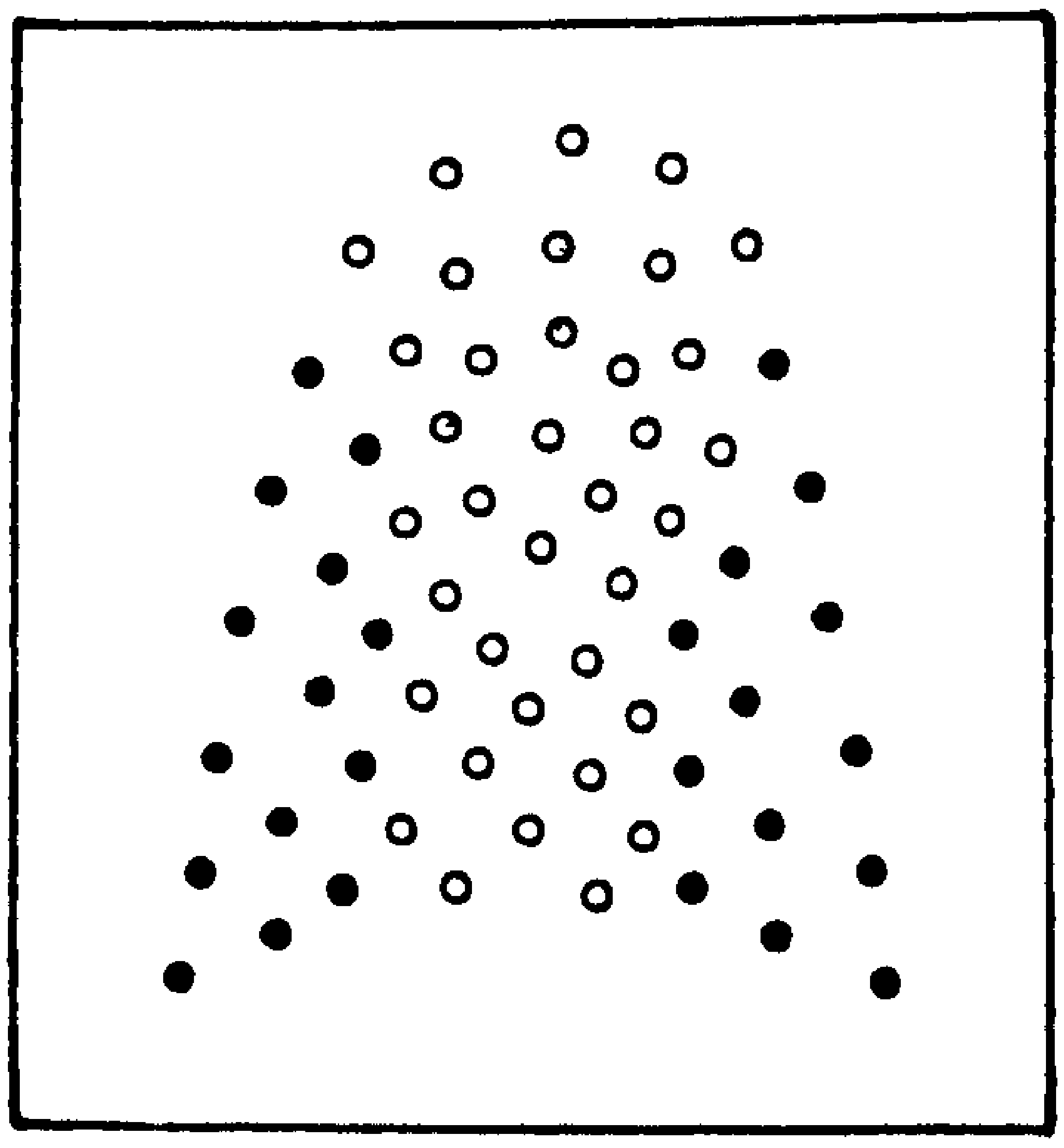
15



16



17



18

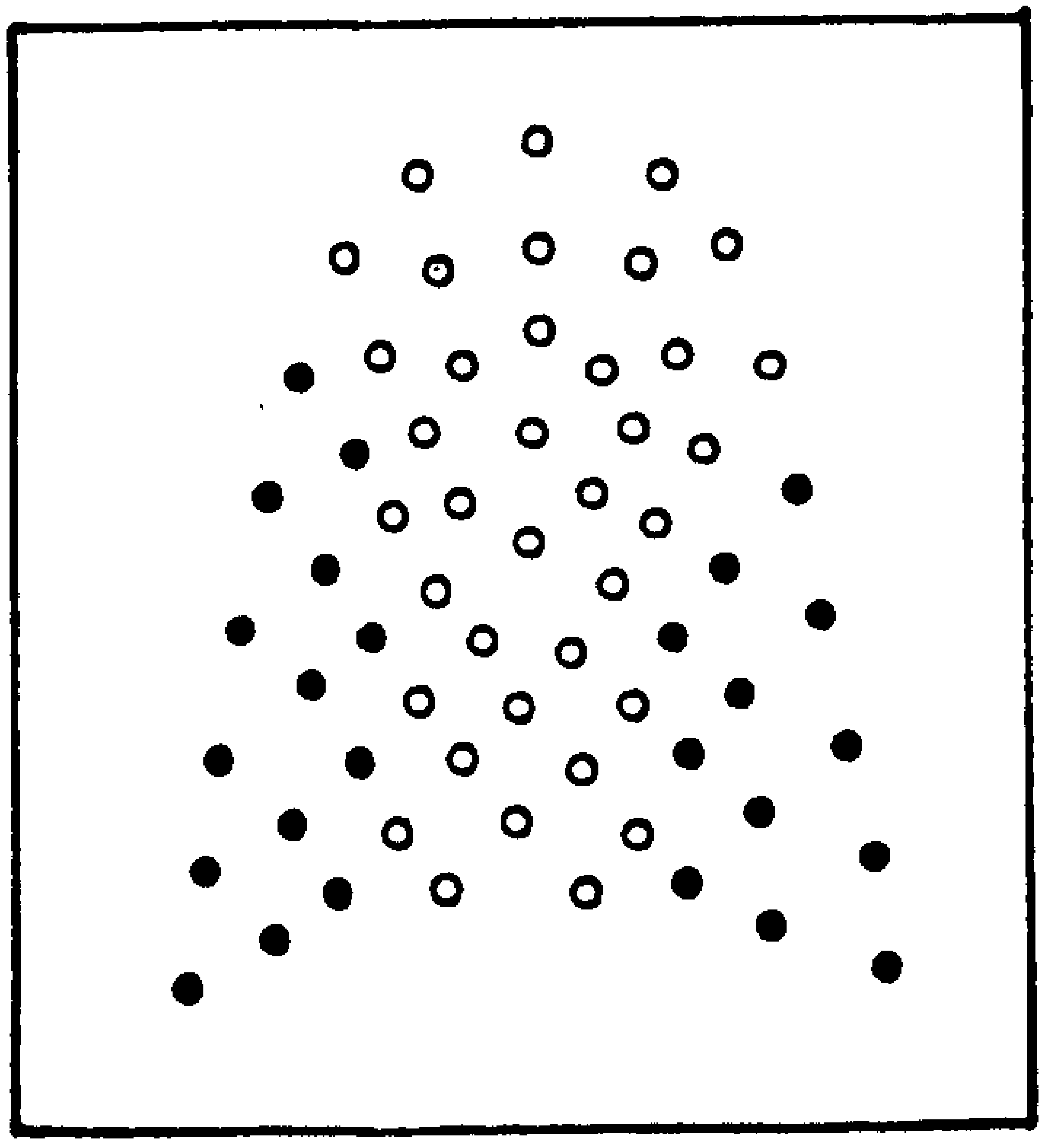


Fig. (5.4) cont'd

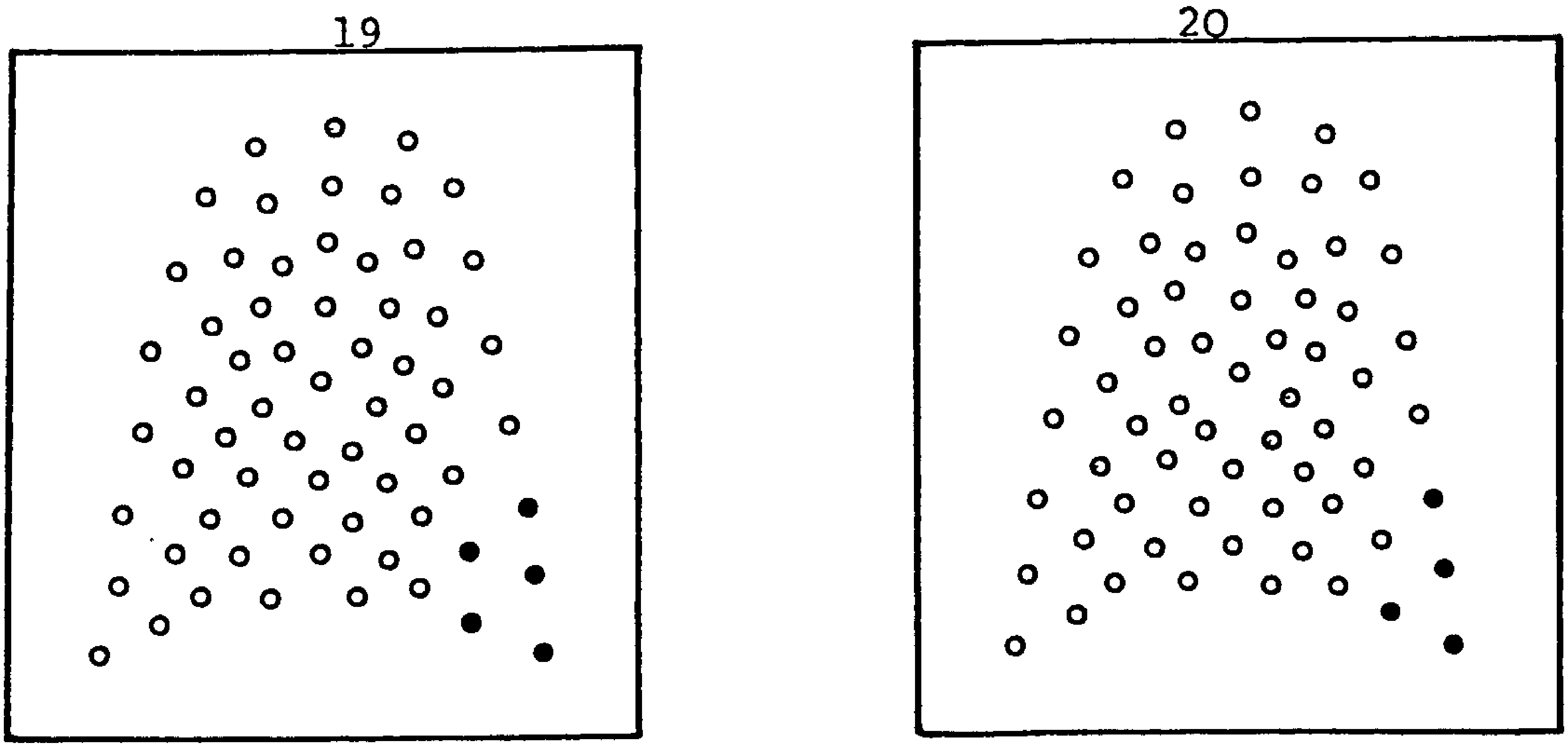


Fig. (5.5) Diagrams showing EPG patterns for the geminate fricative [ss] in the environment of /ξassa/.

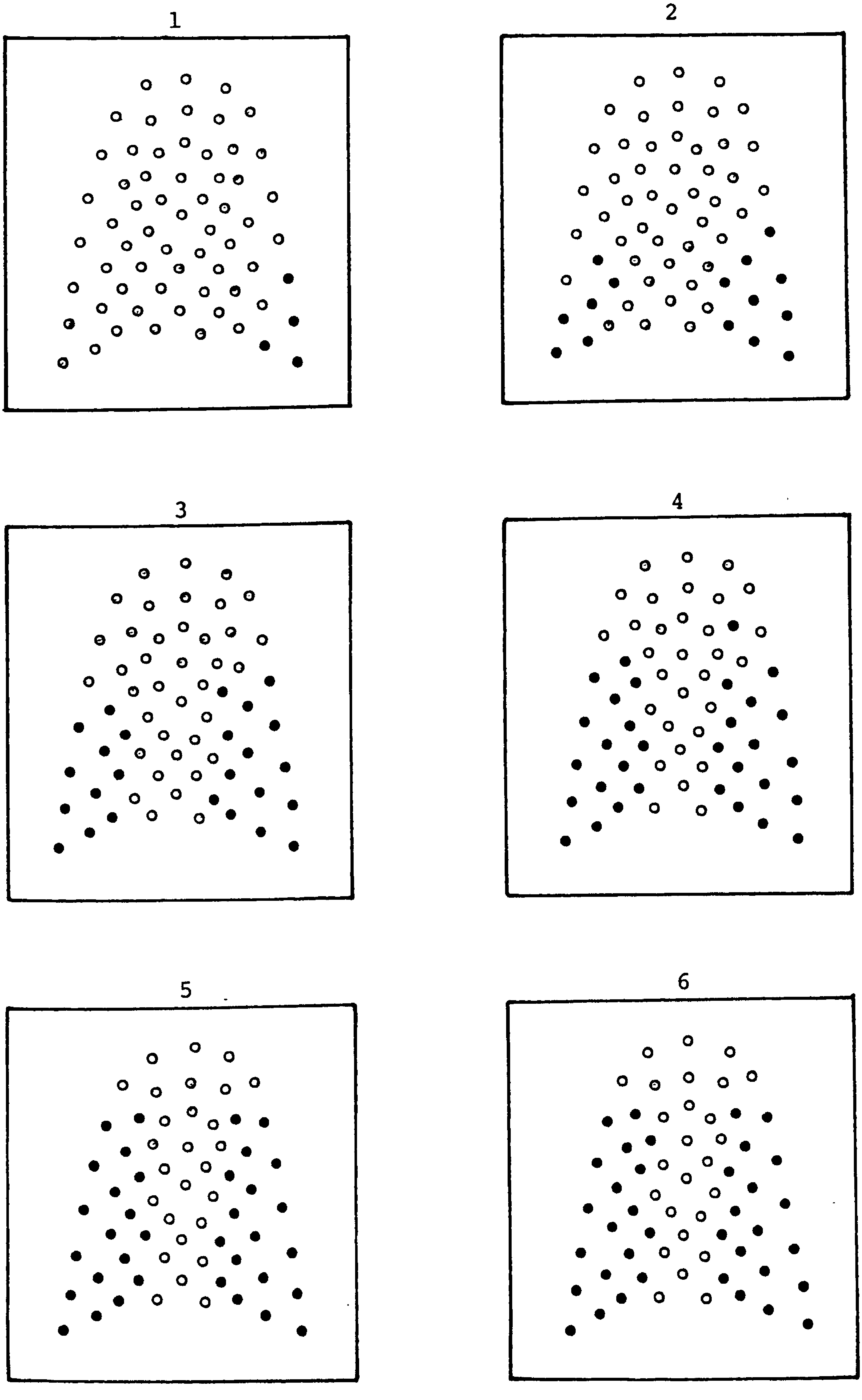


Fig. (5.5) cont'd

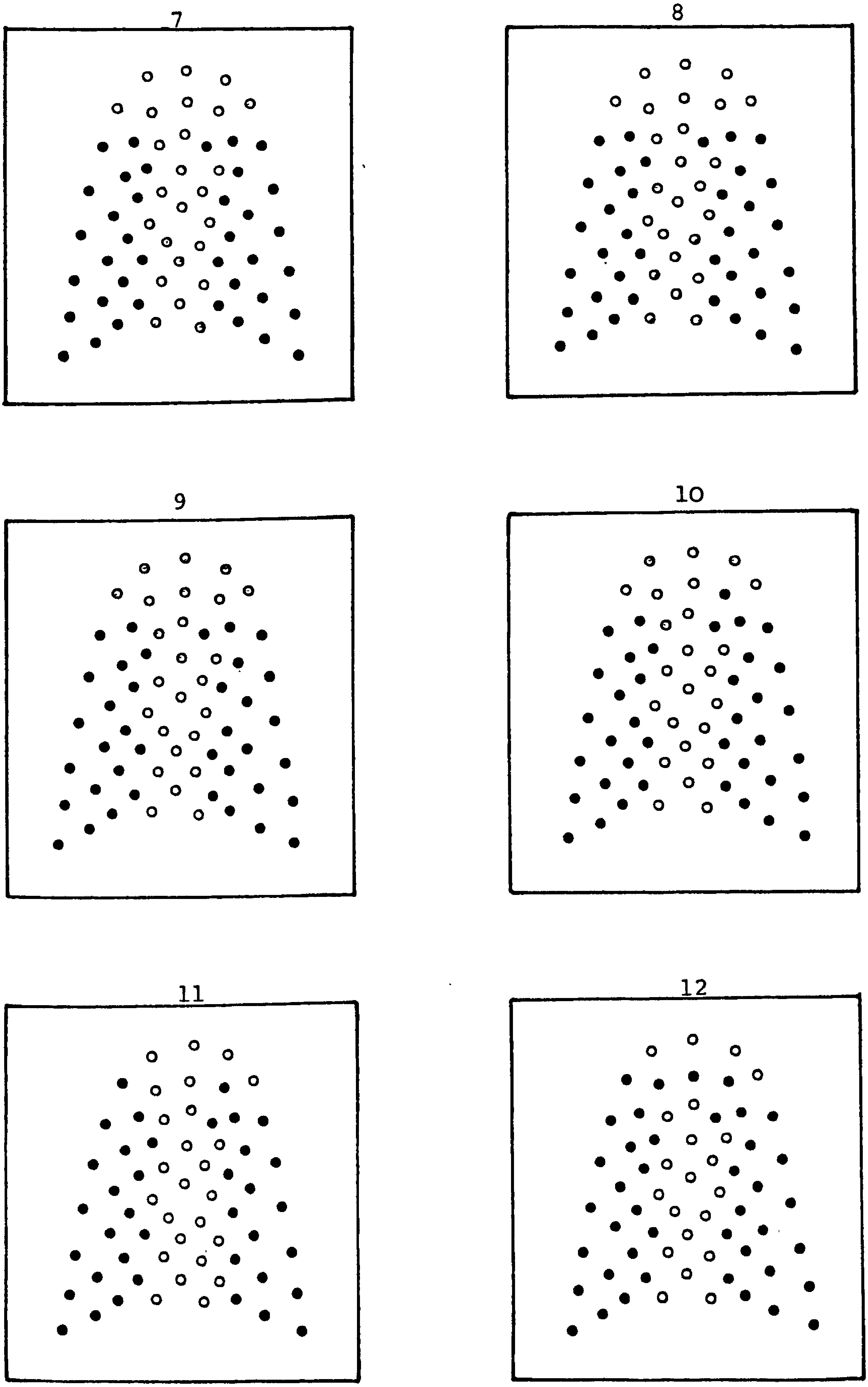
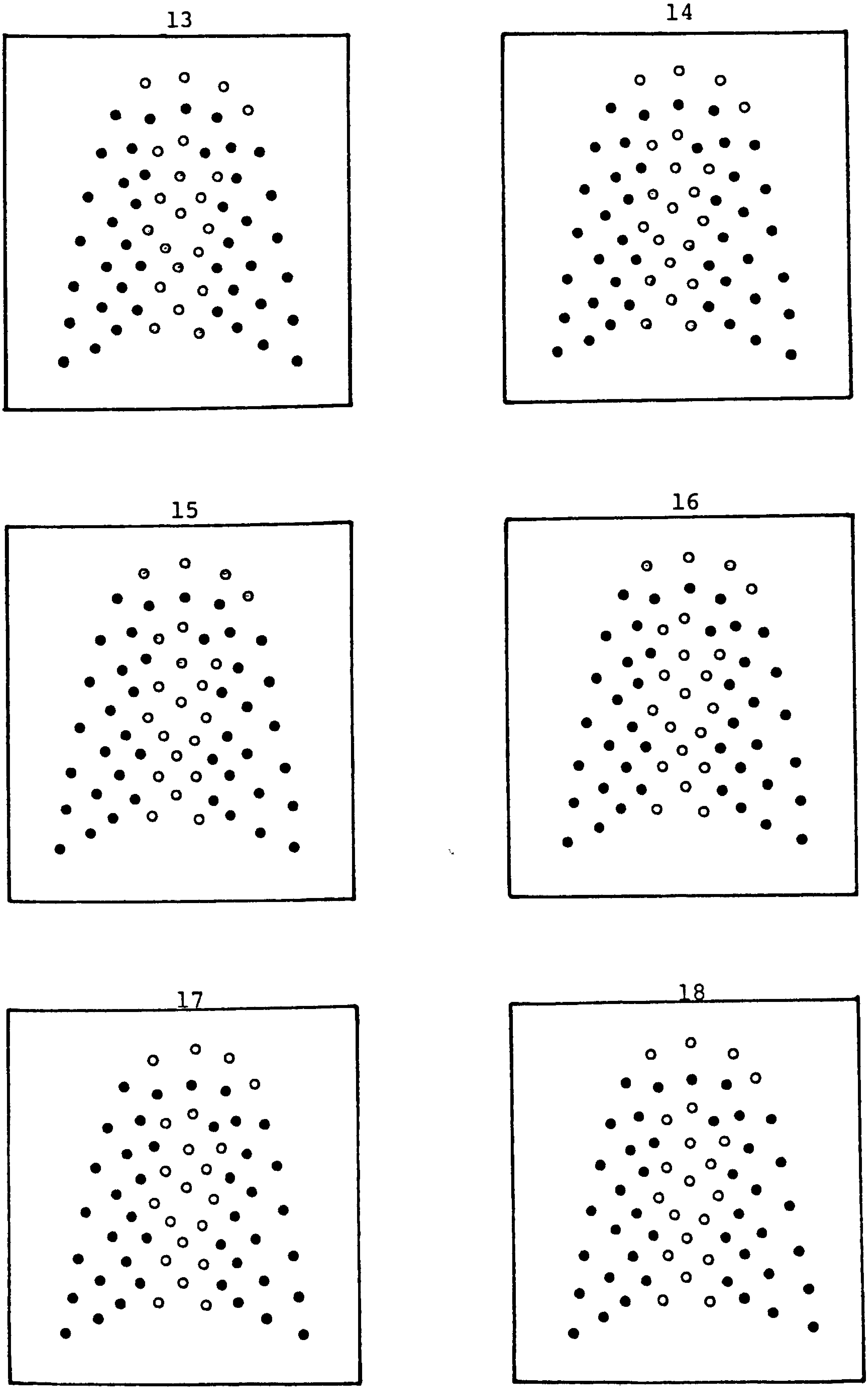


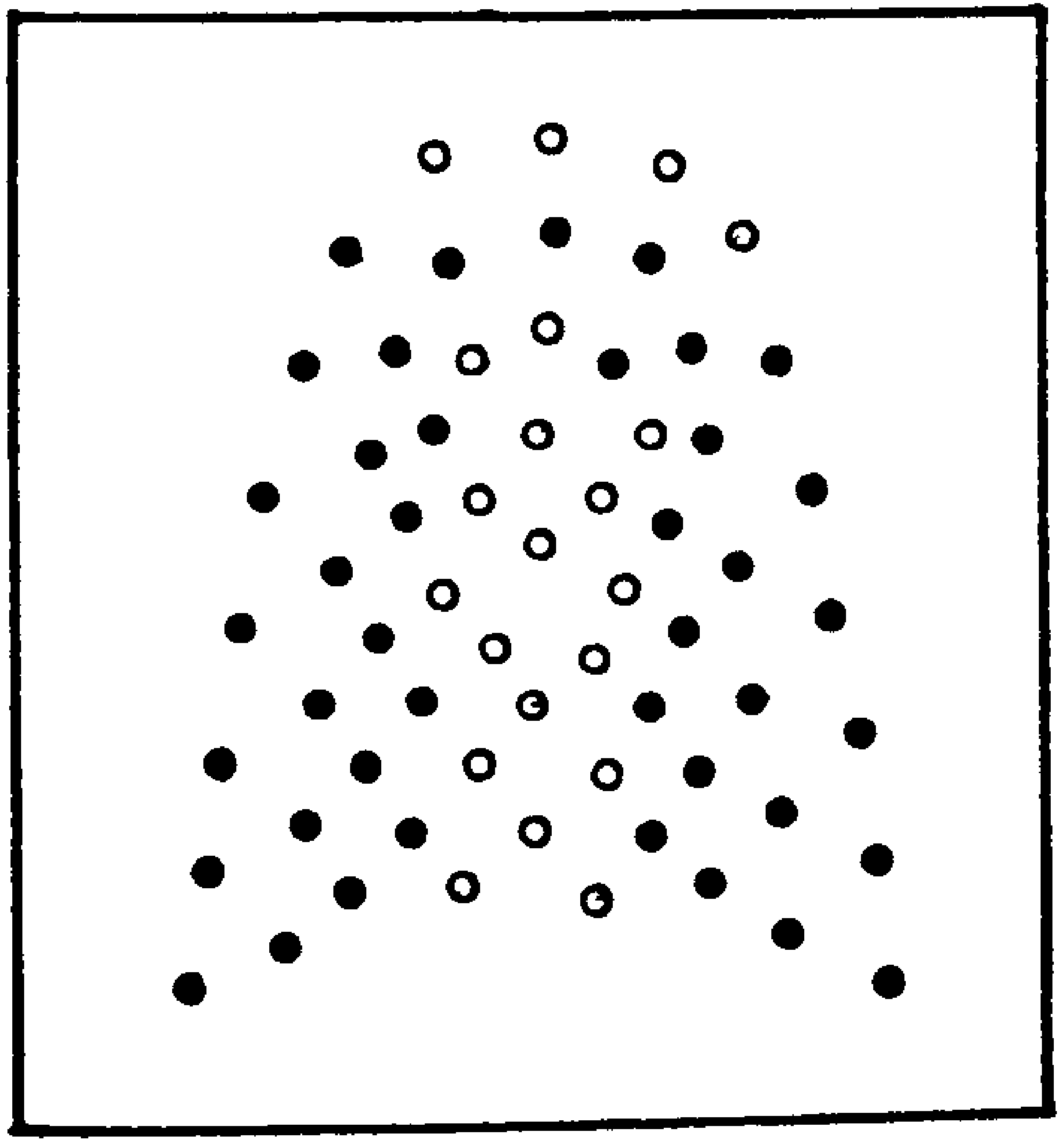
Fig. (5.5) cont'd



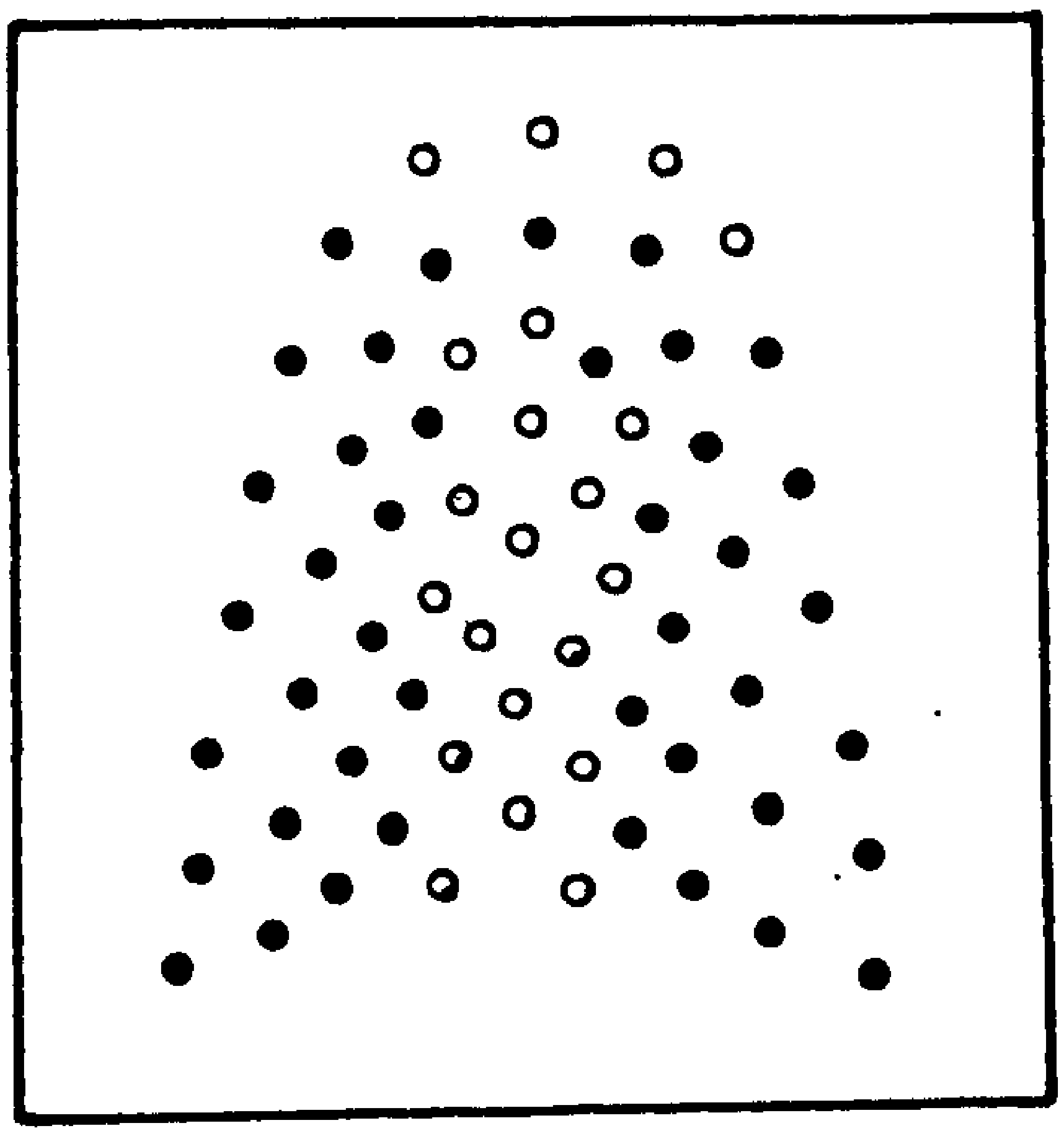
13 14 15 16 17 18

Fig. (5.5) cont'd

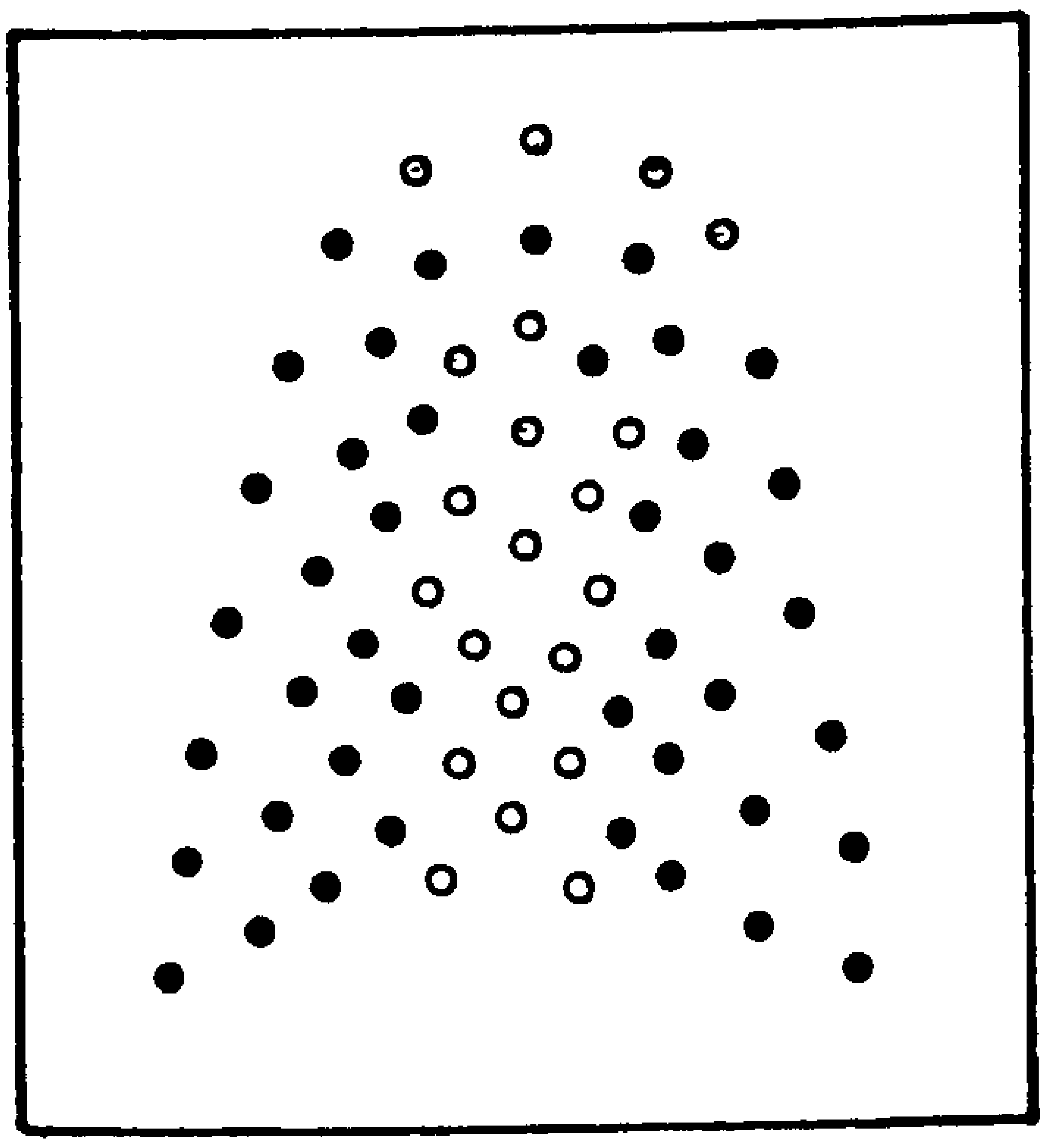
19



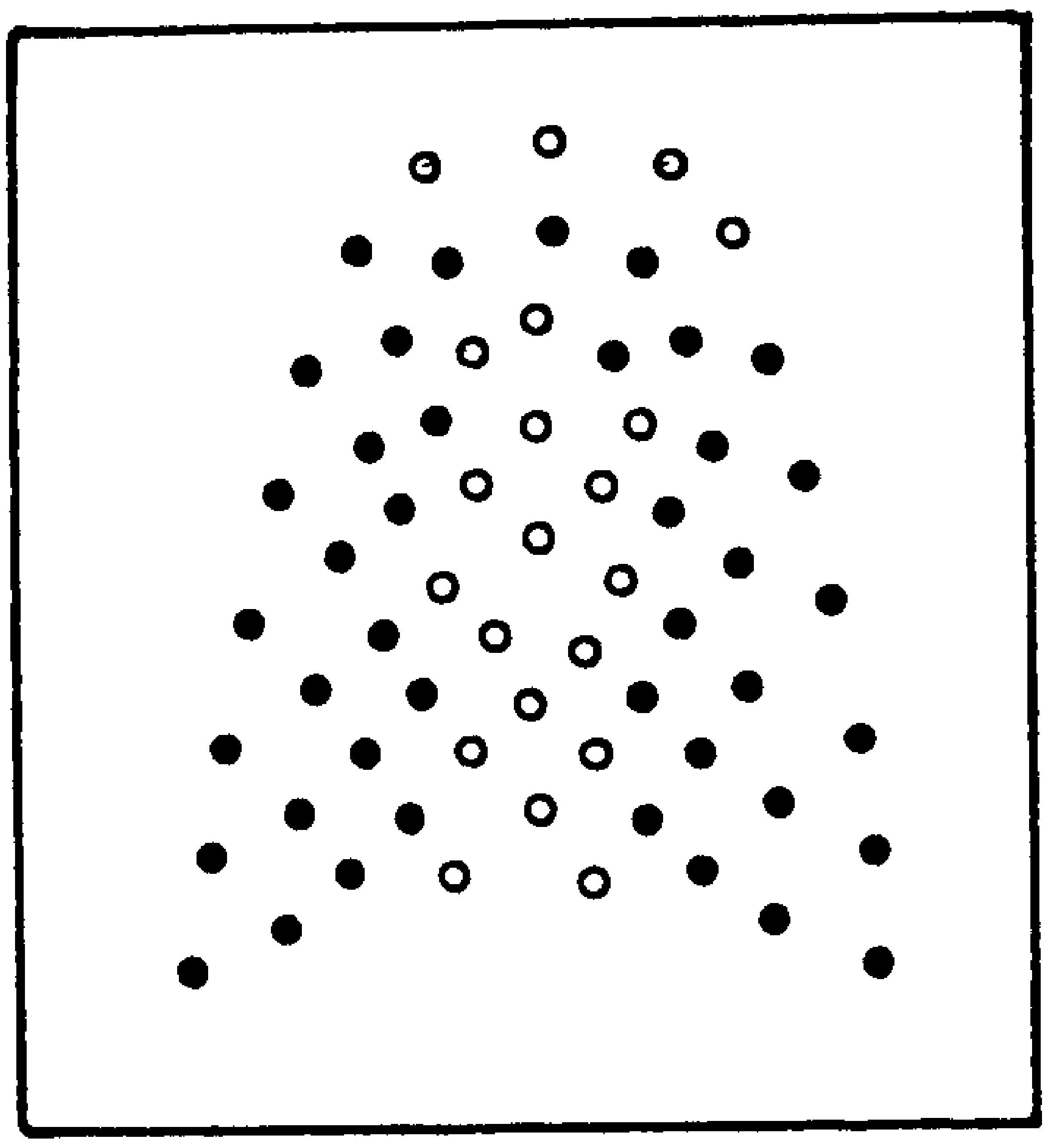
20



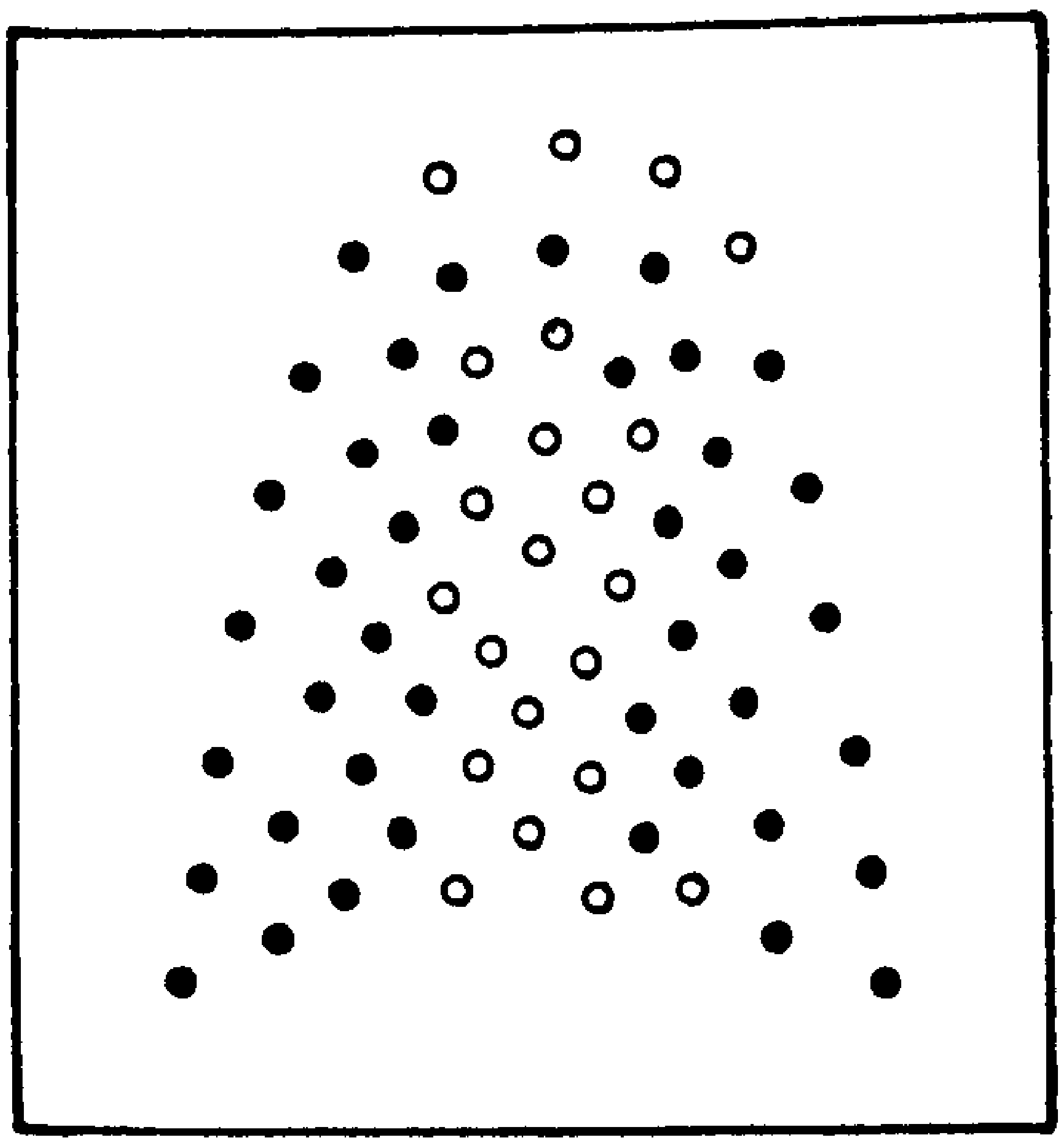
21



22



23



24

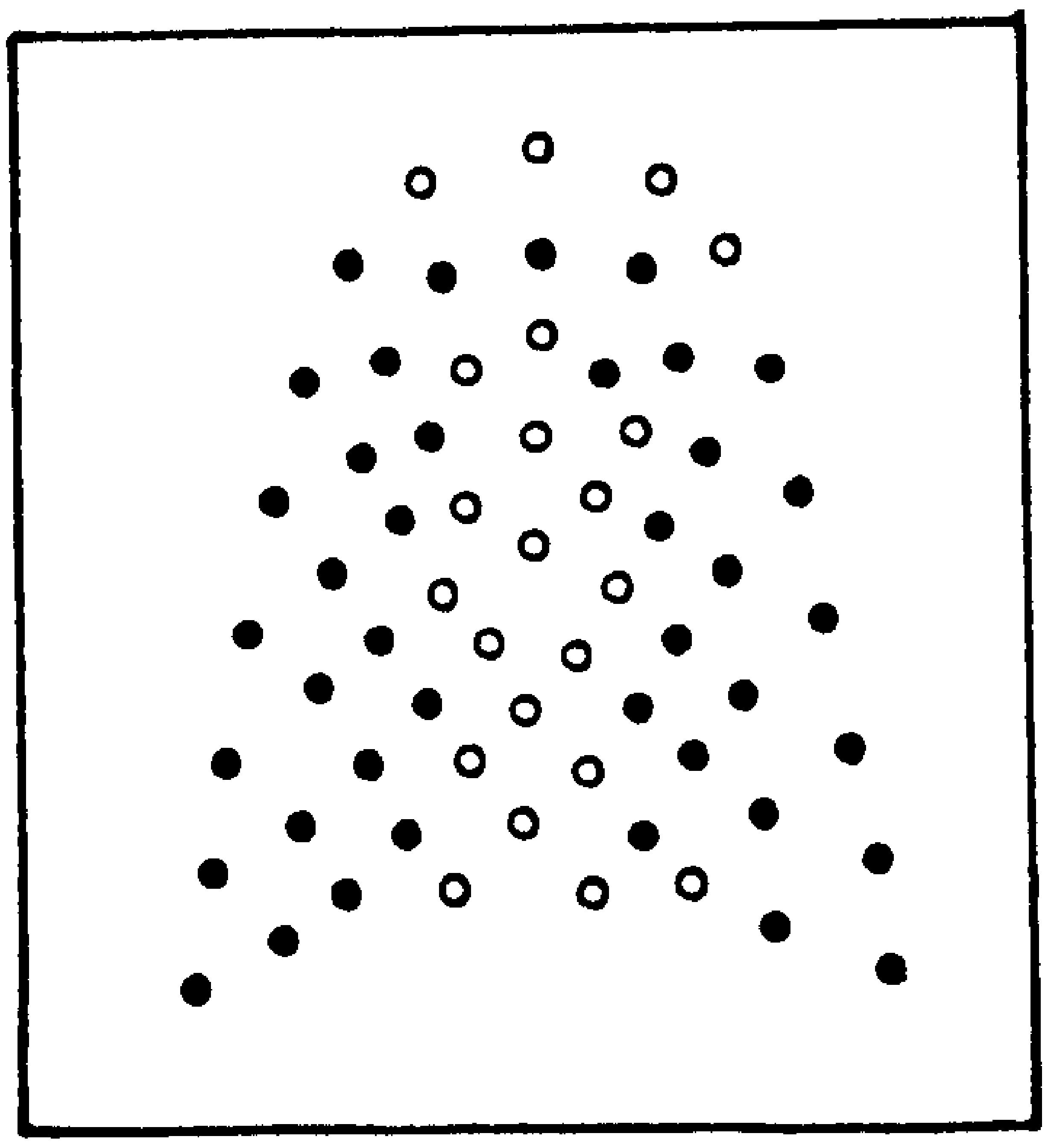
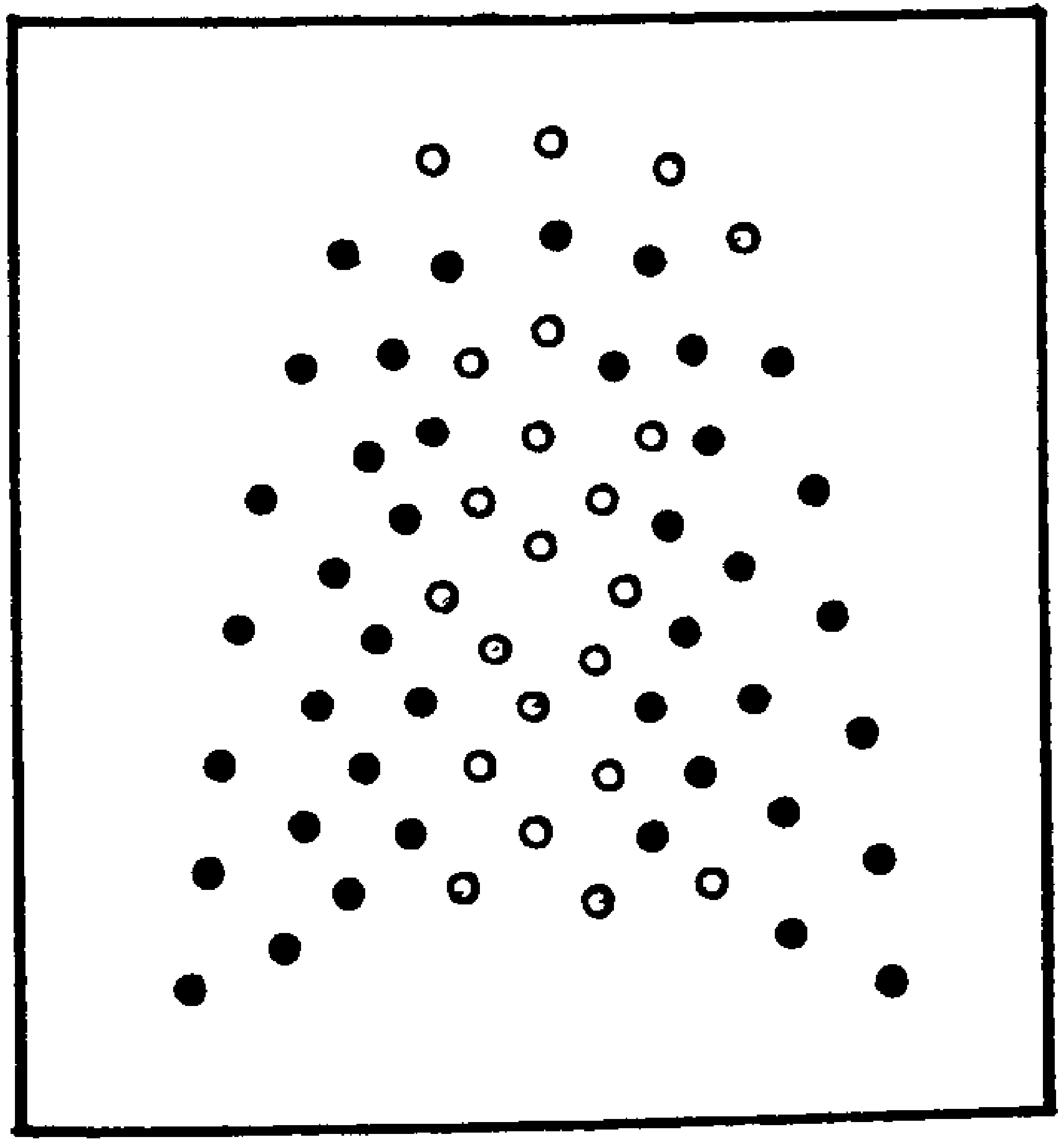
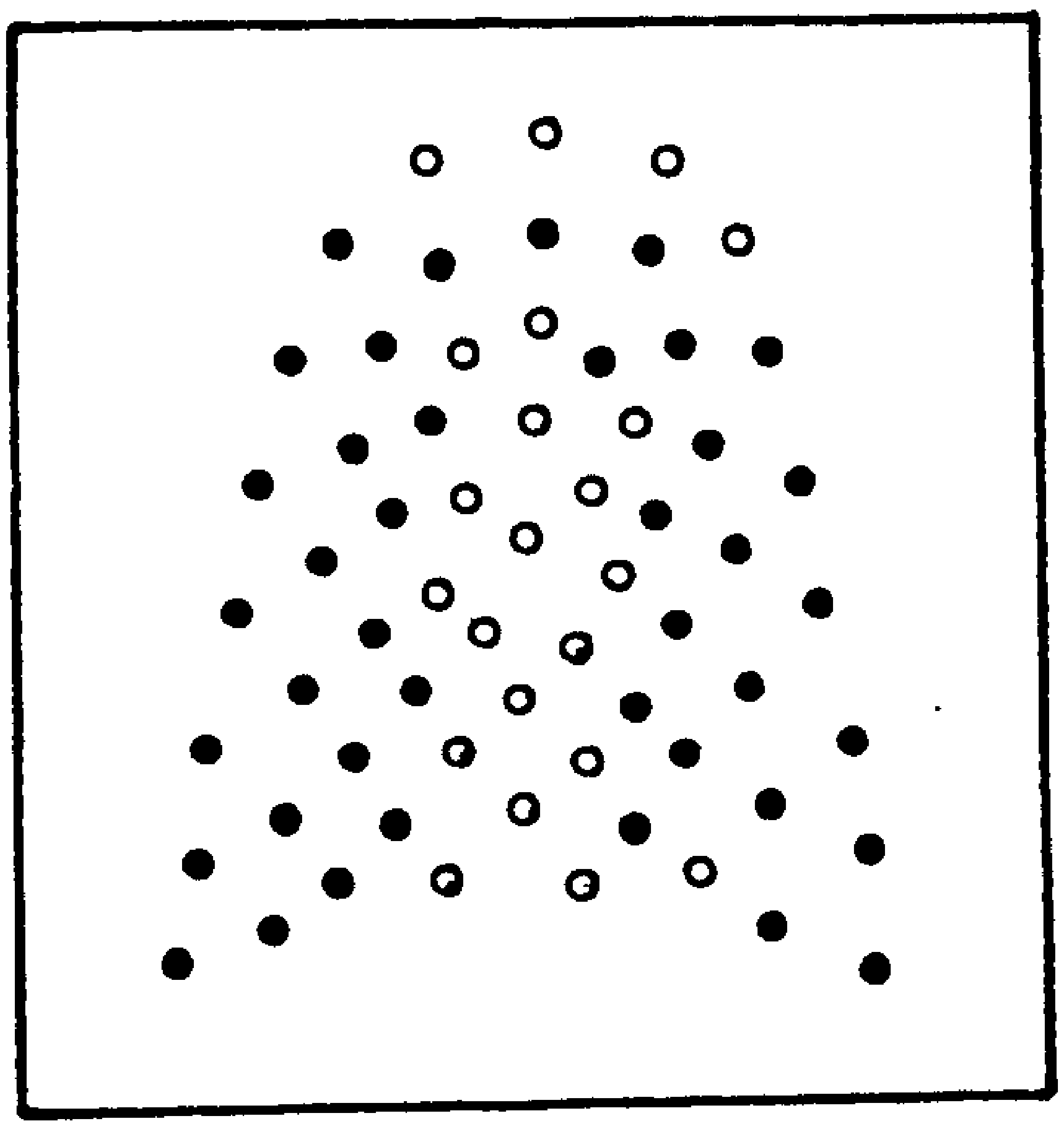


Fig. (5.5) cont'd

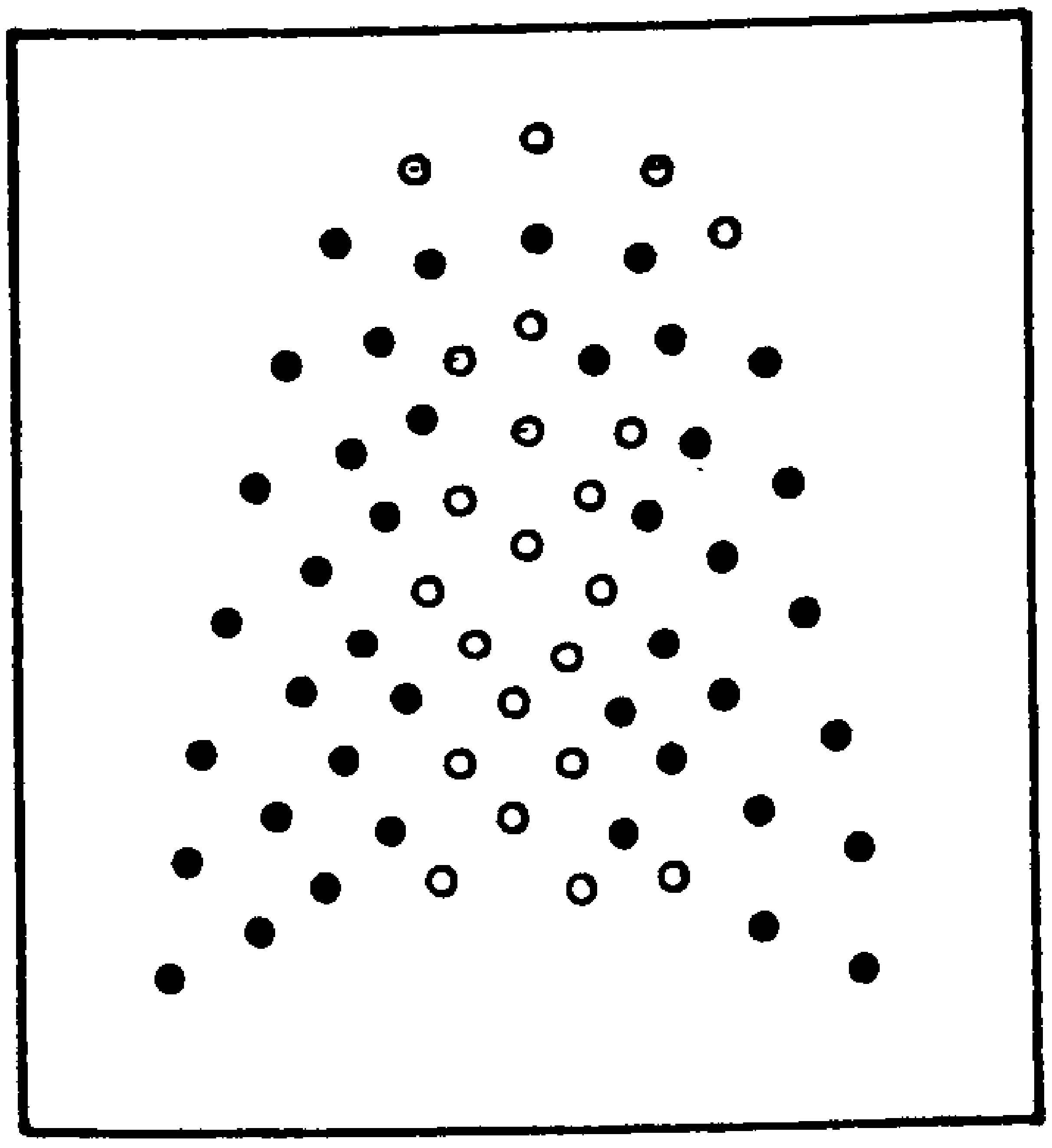
25



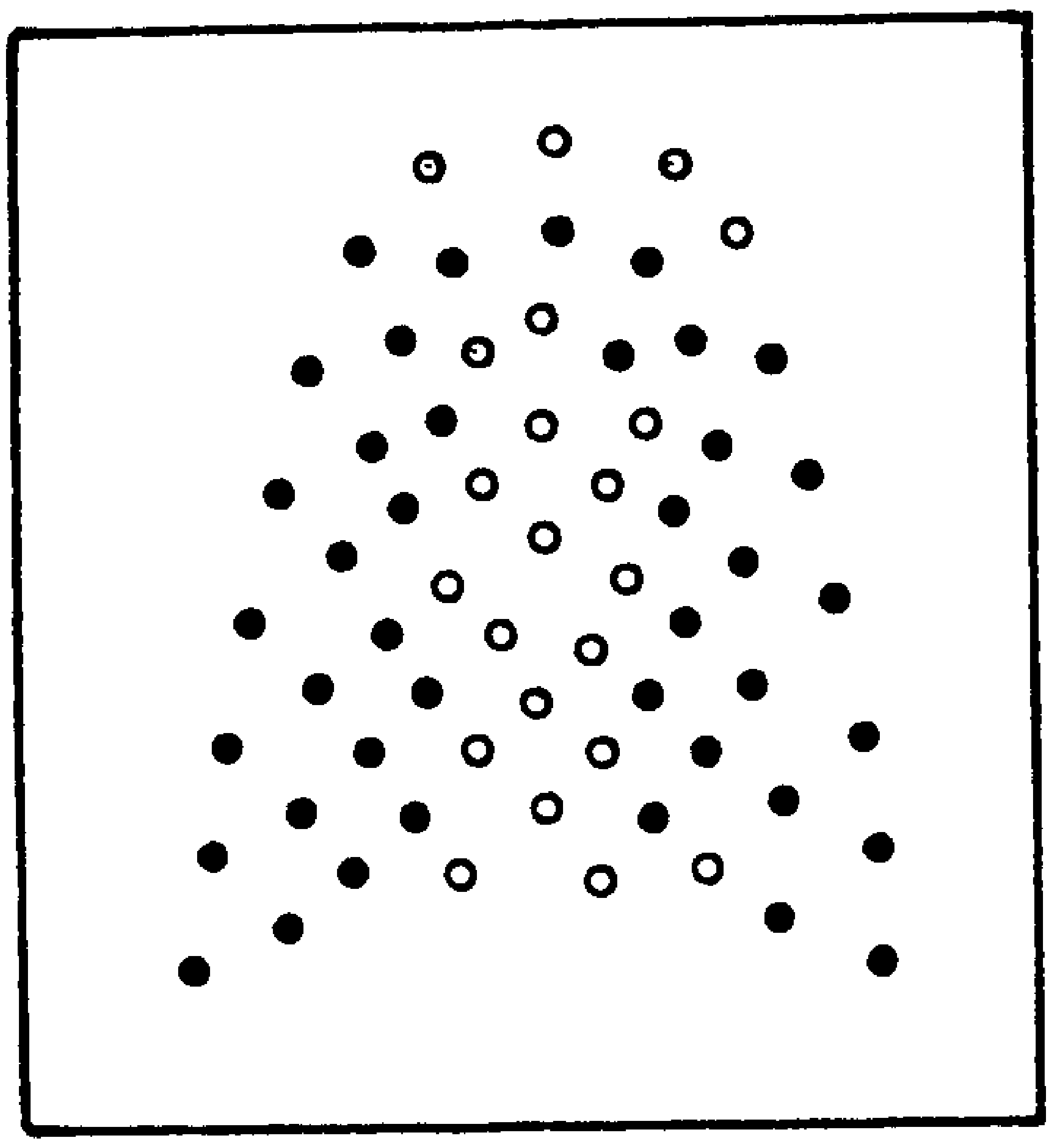
26



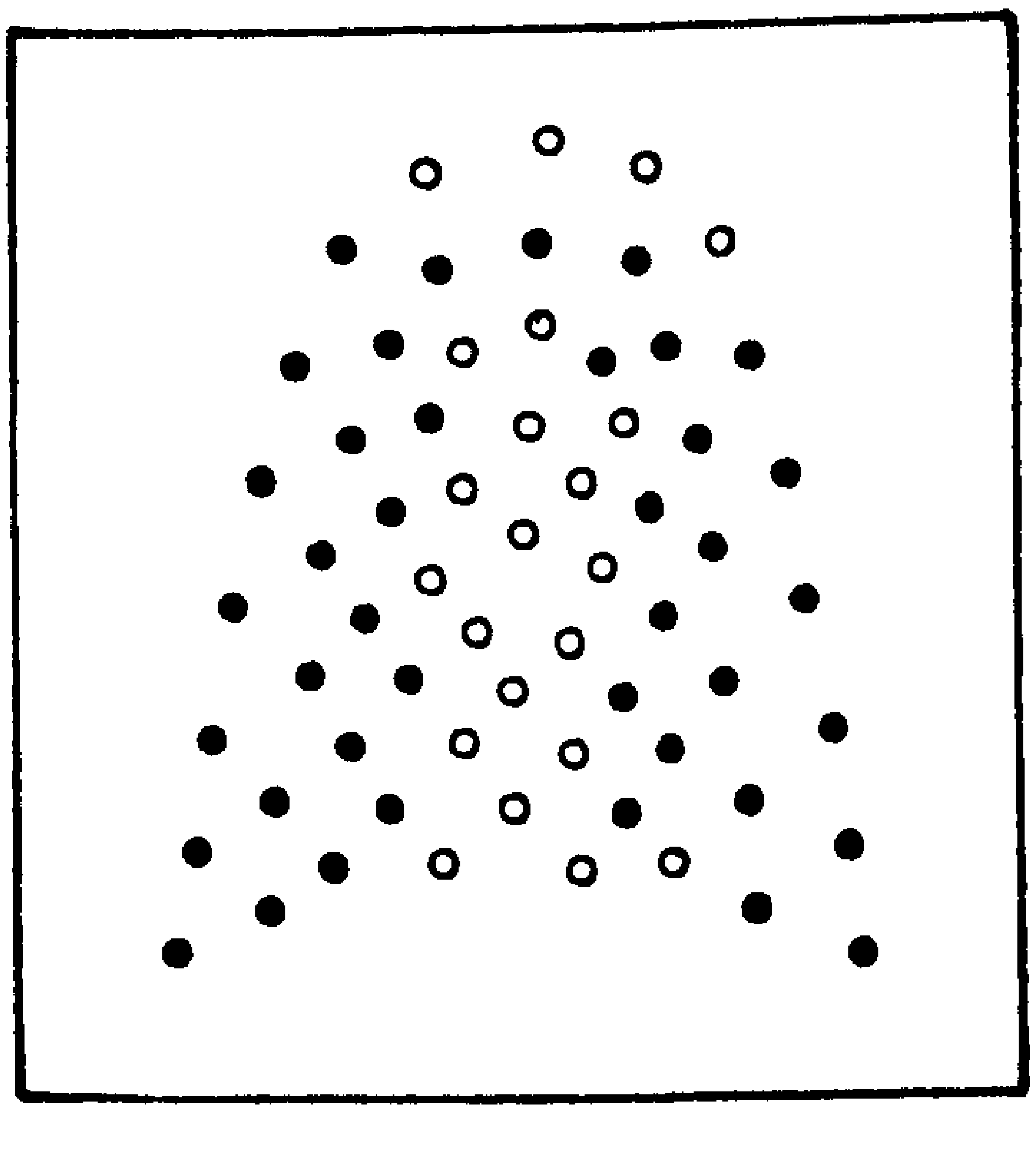
27



28



29



30

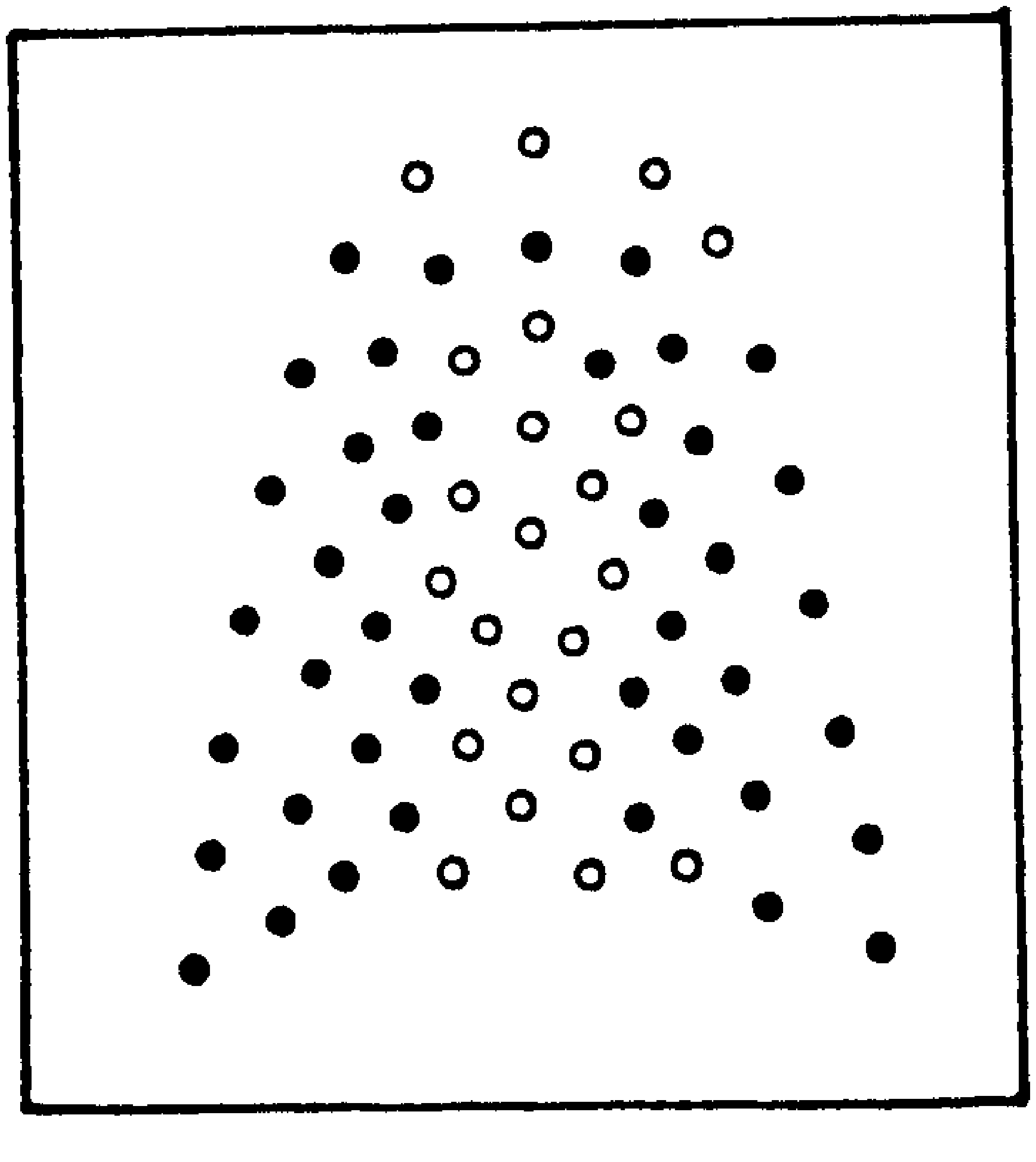
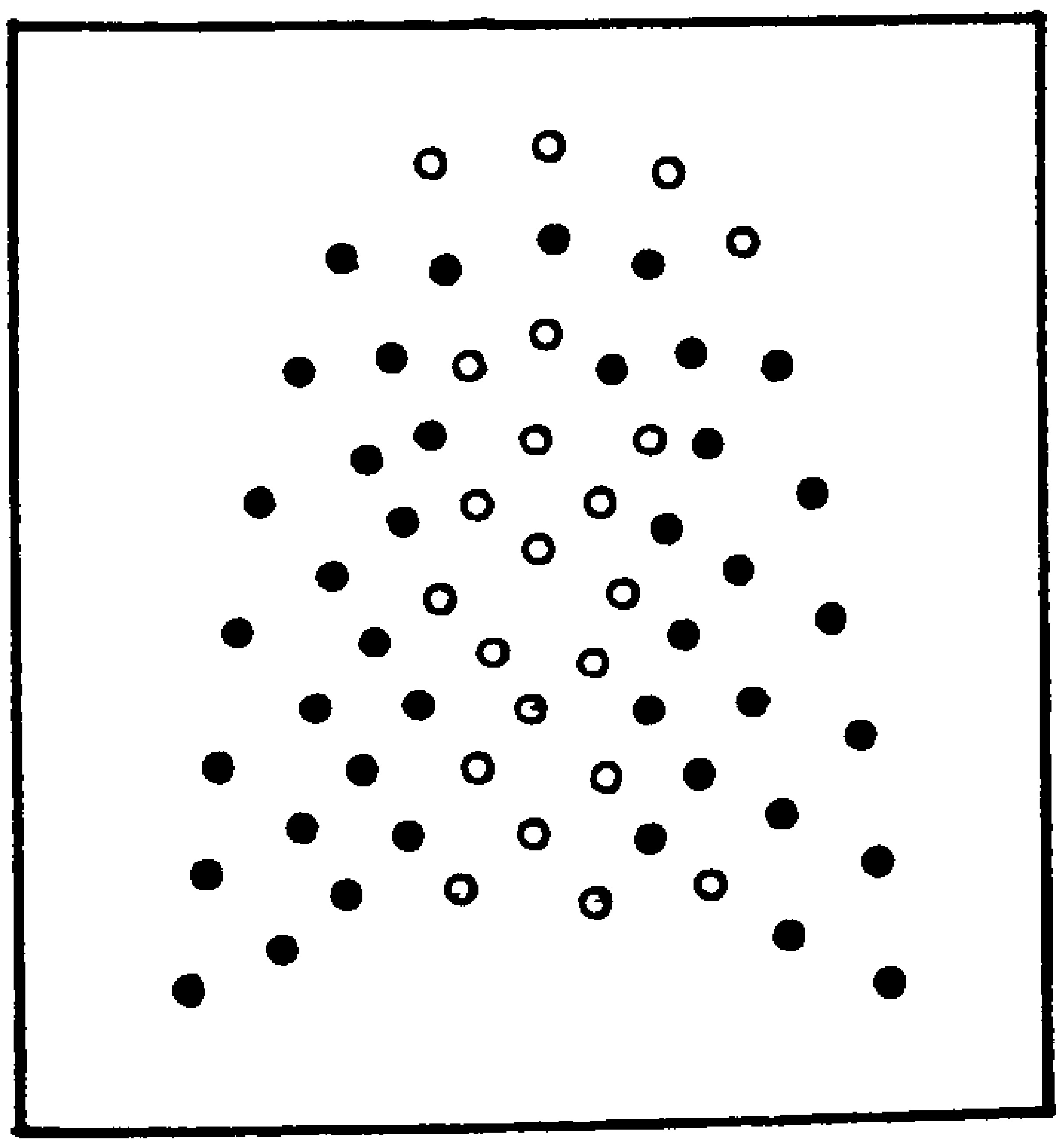
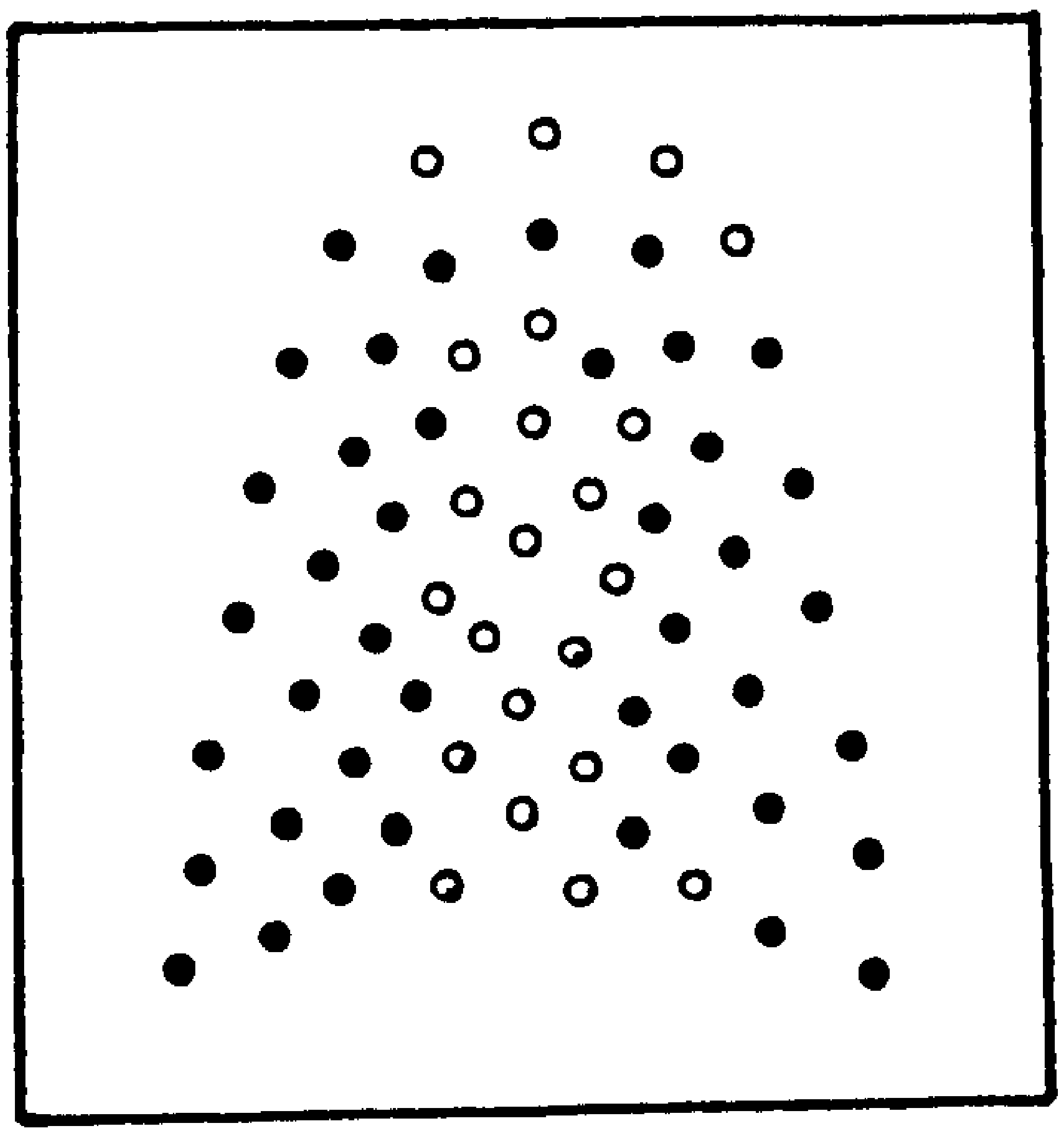


Fig. (5.5) cont'd

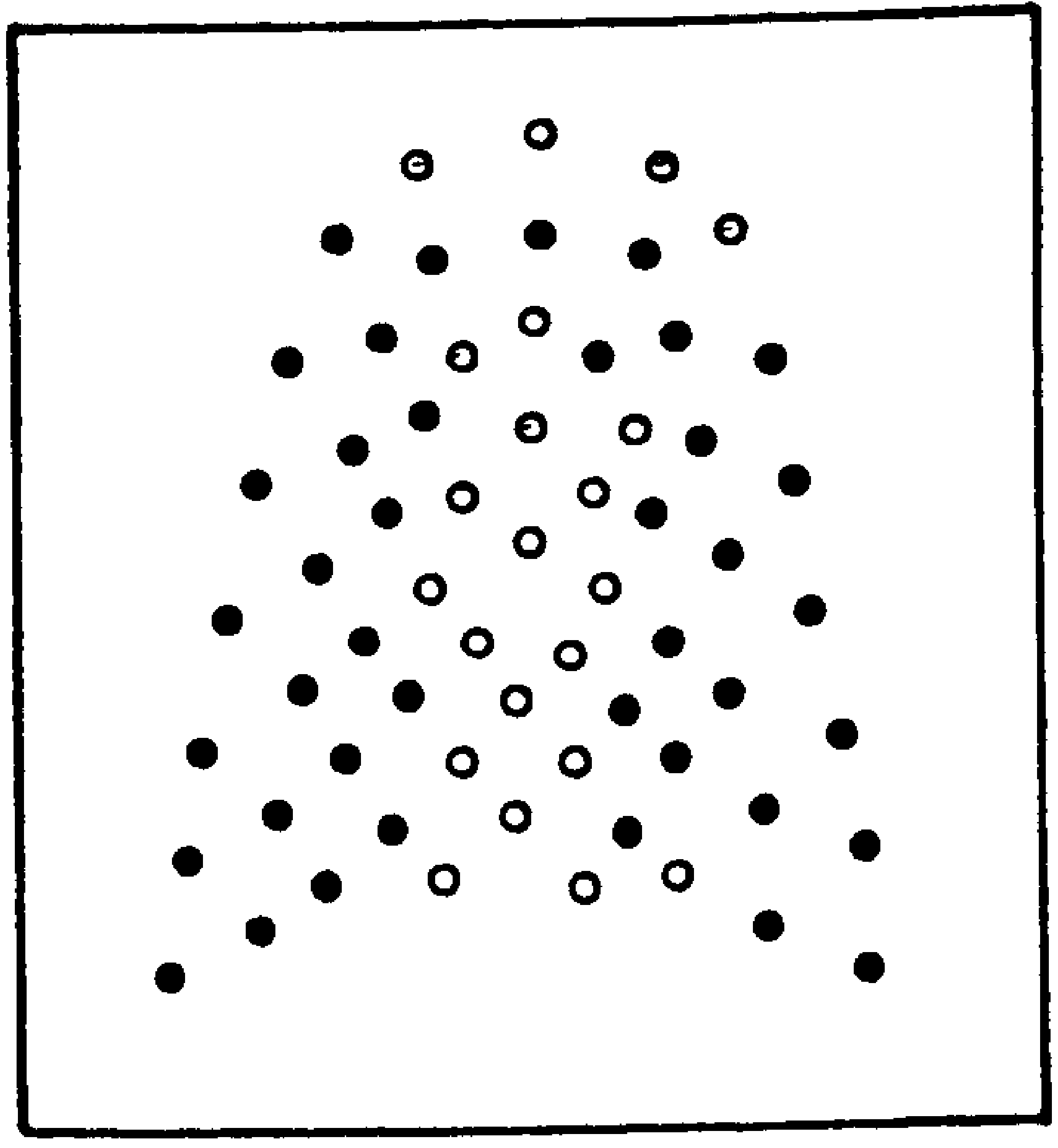
31



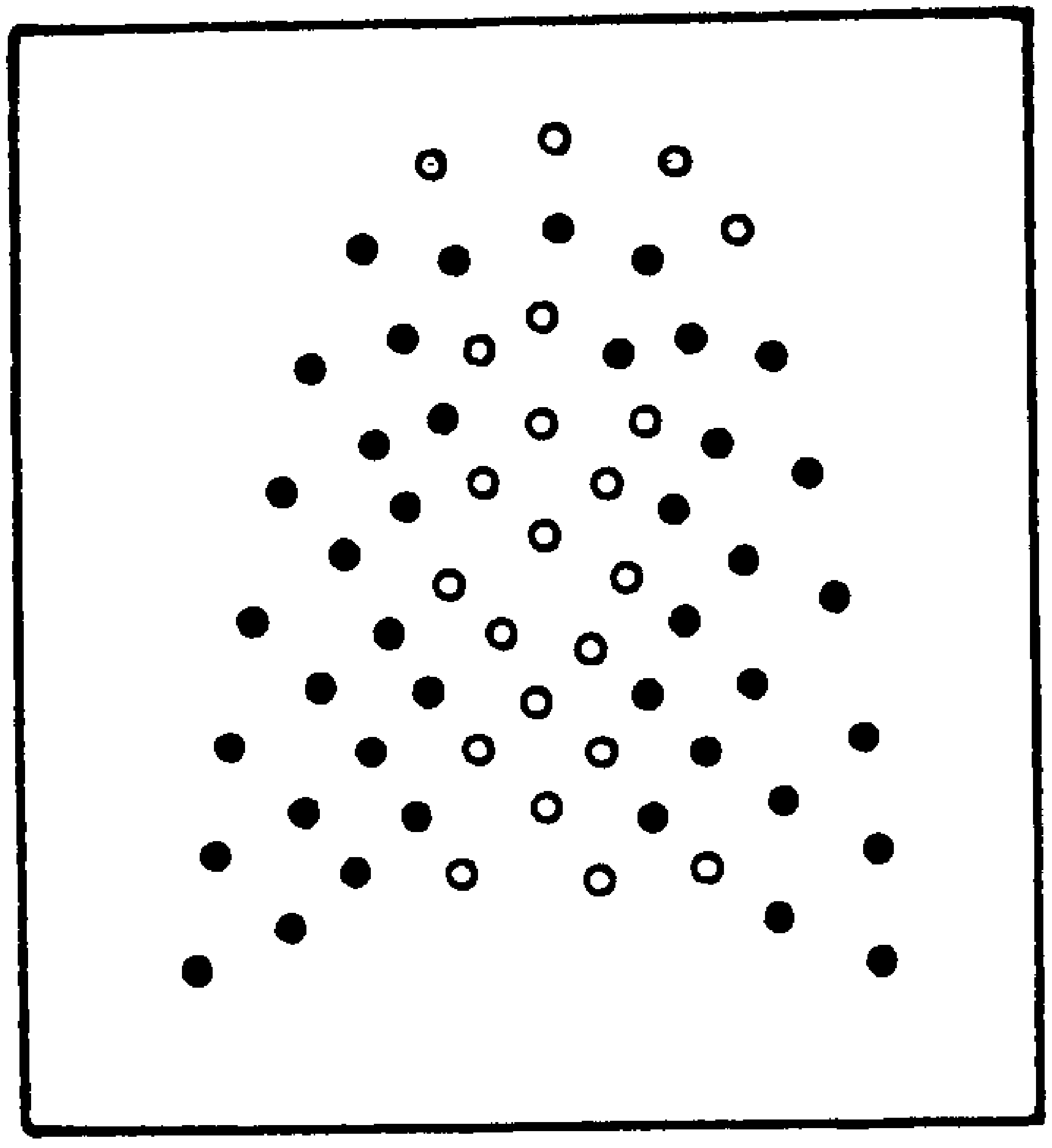
32



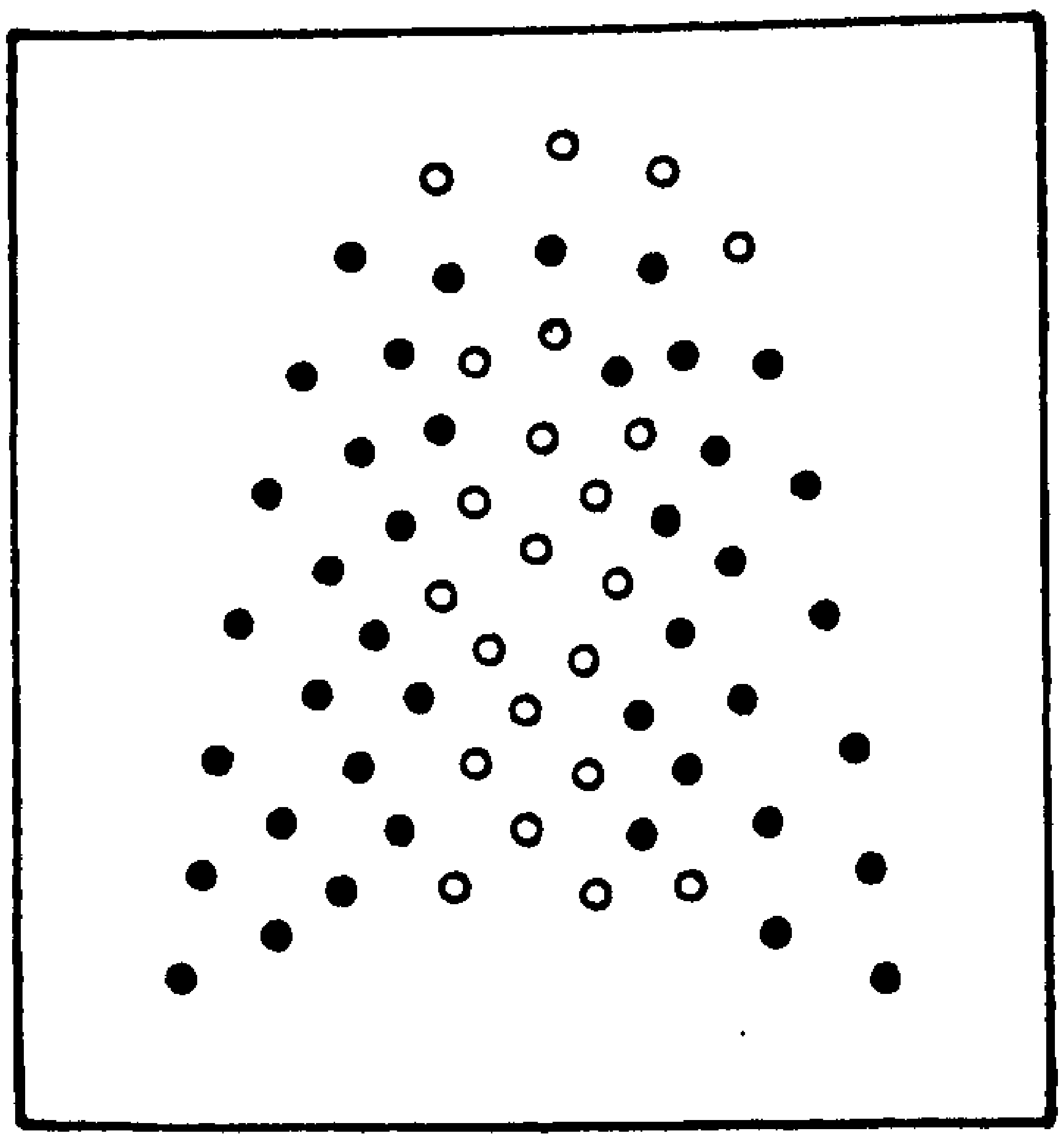
33



34



35



36

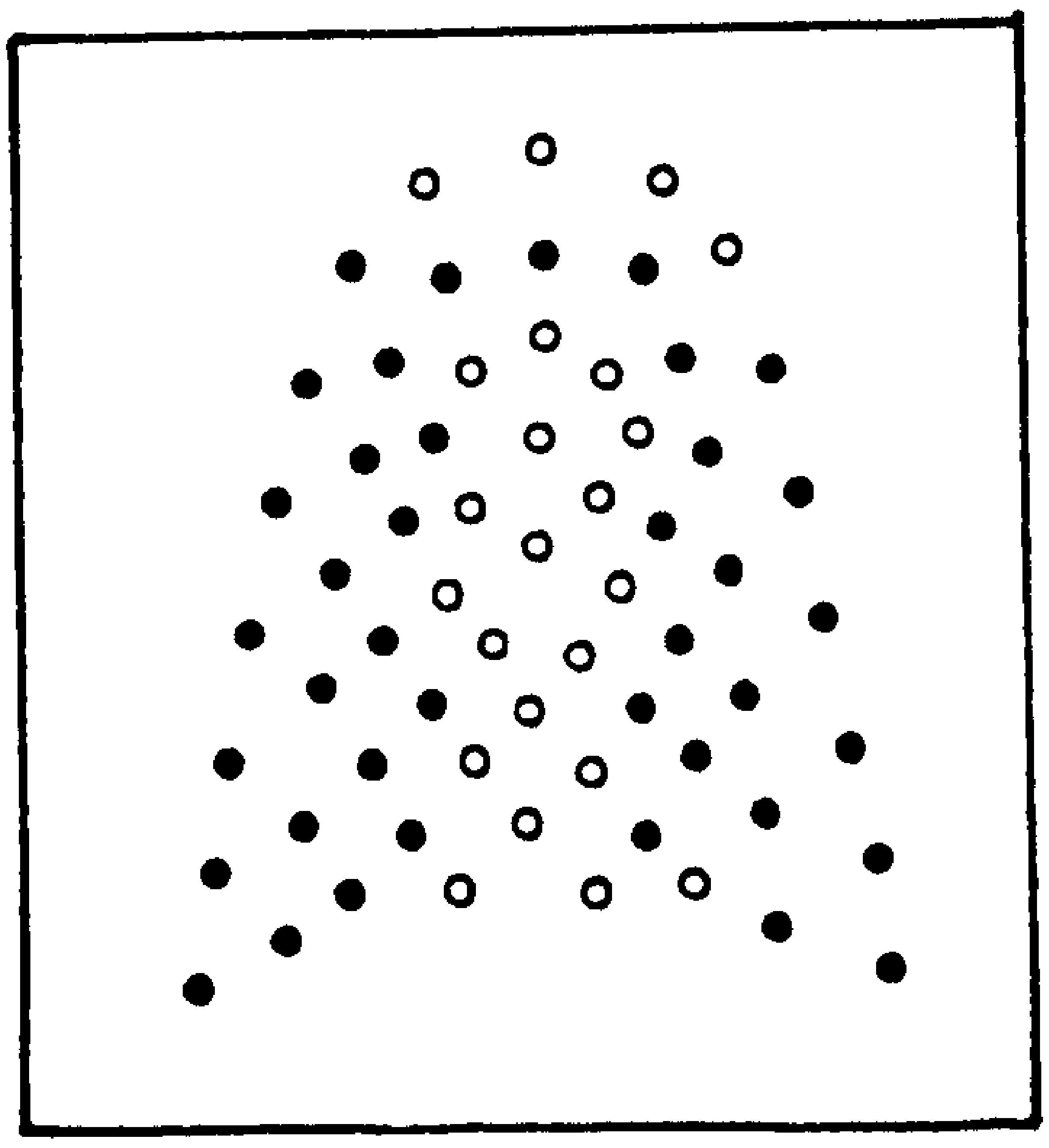
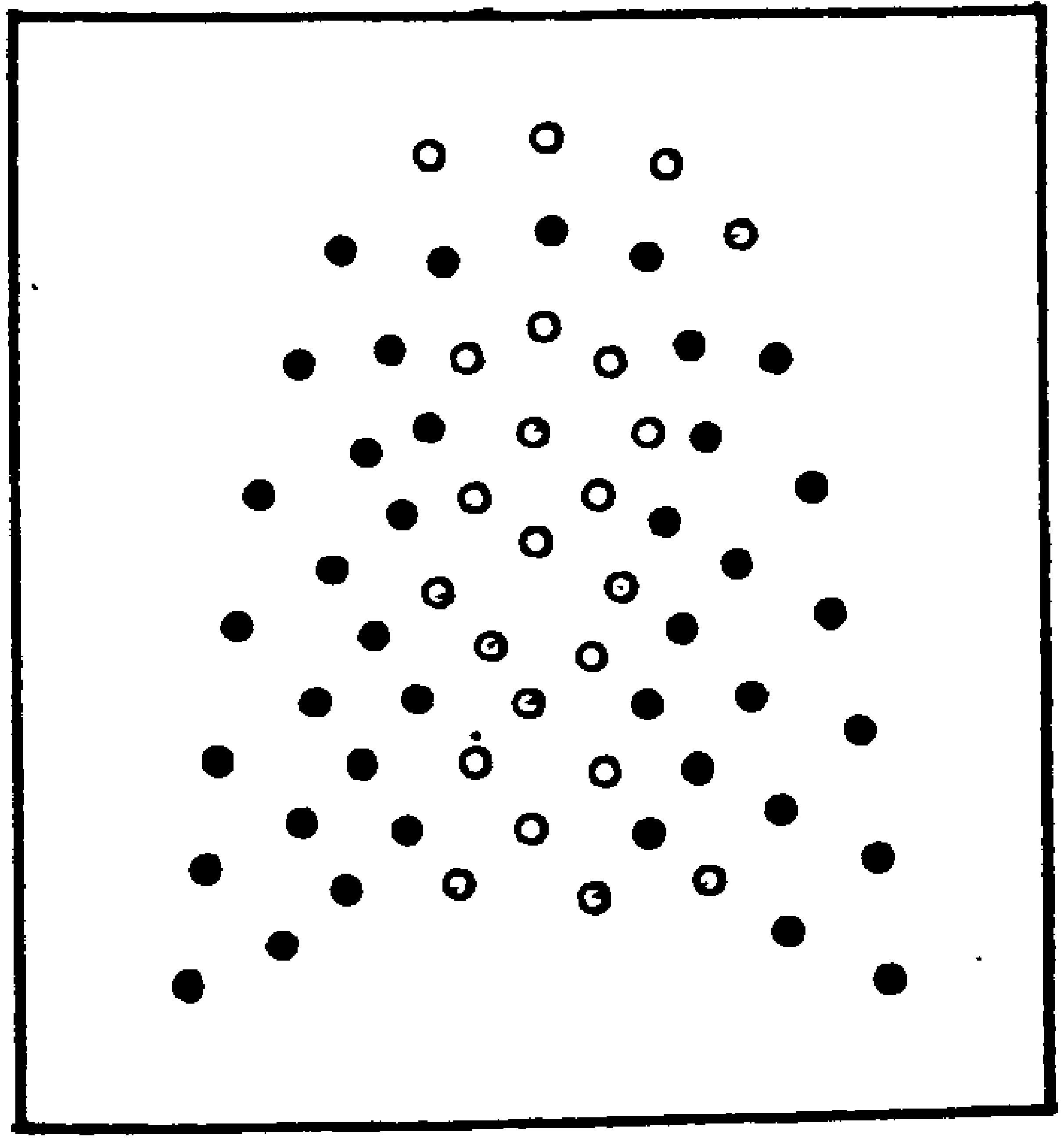
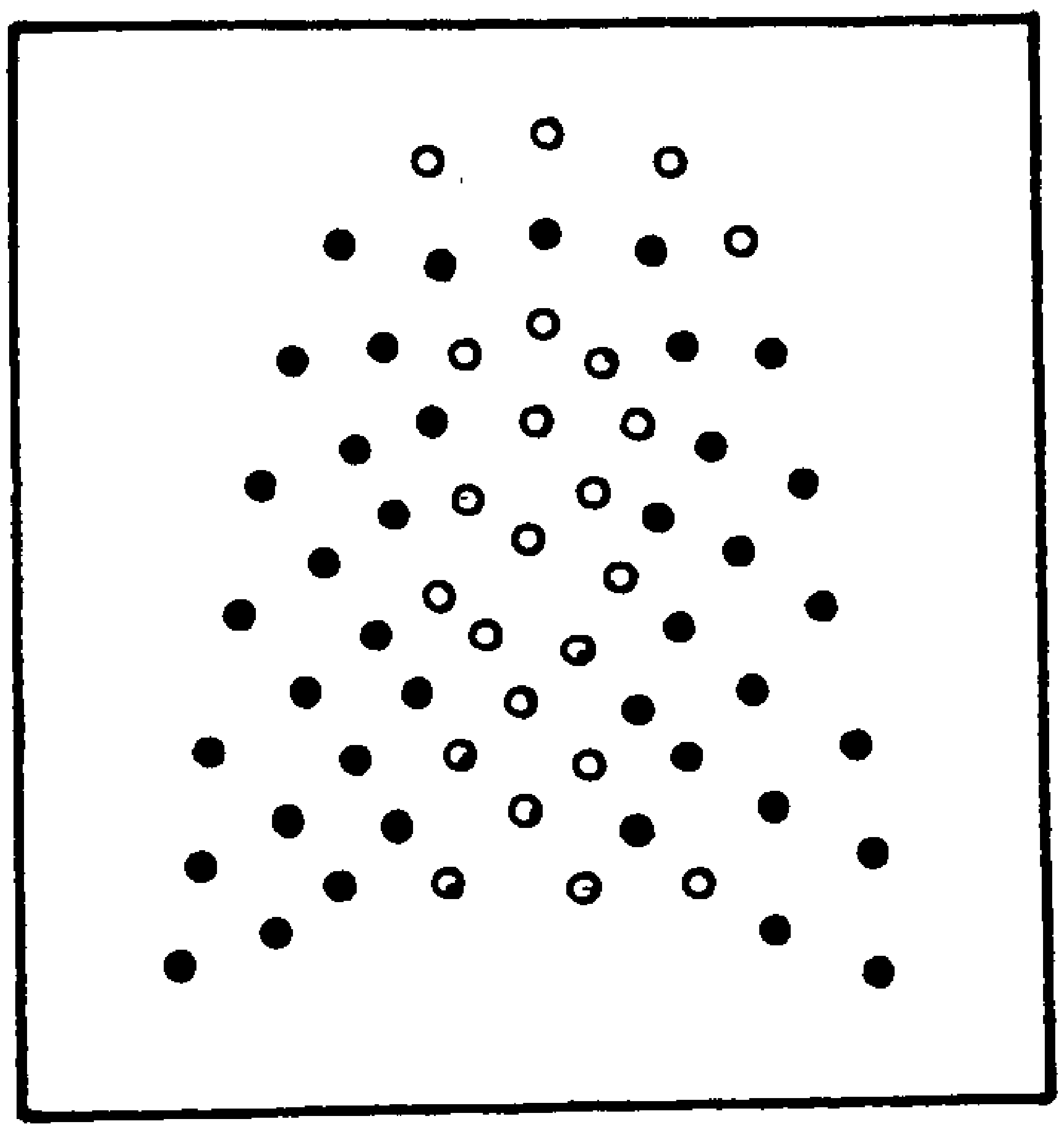


Fig. (5.5) cont'd

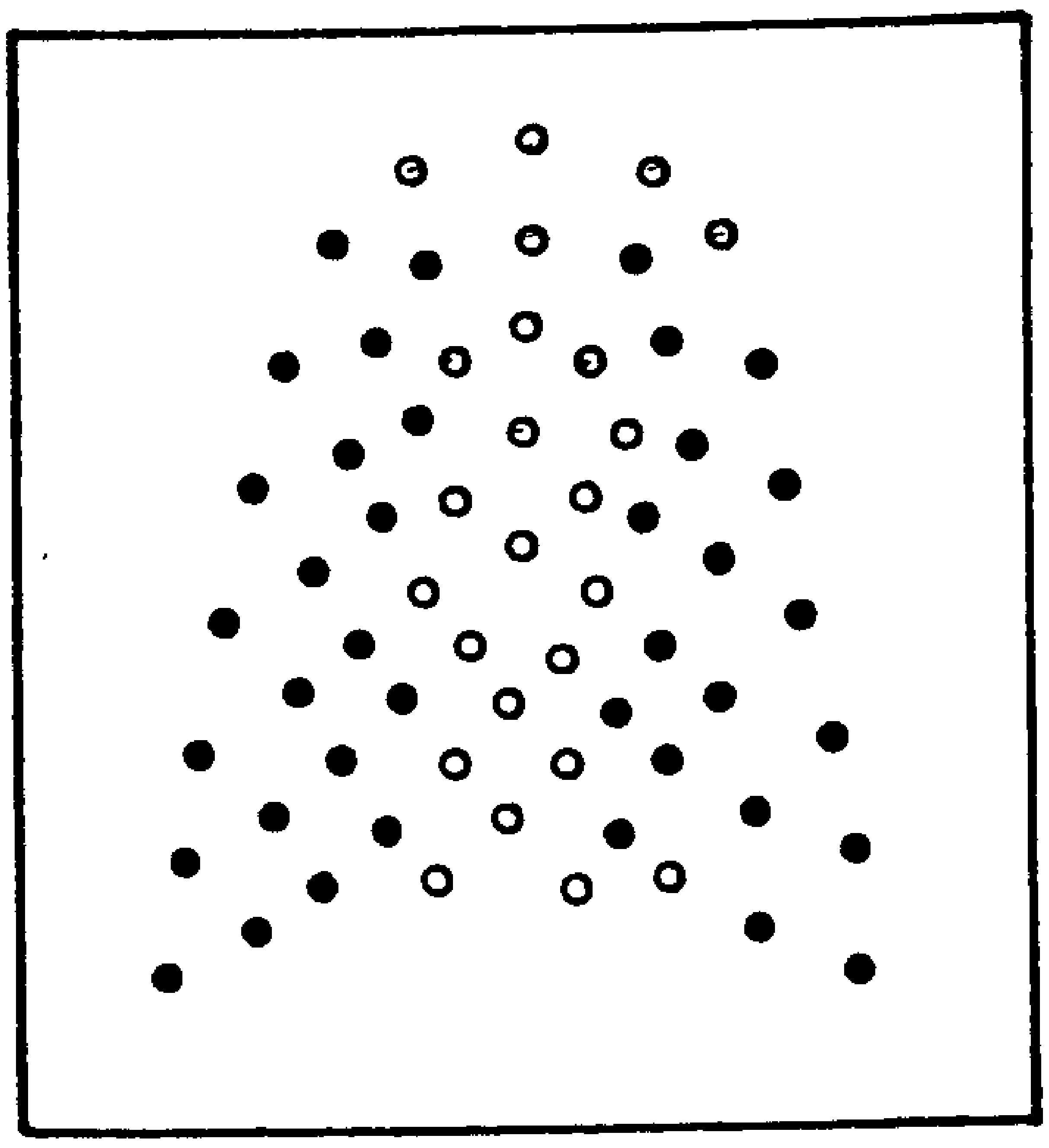
37



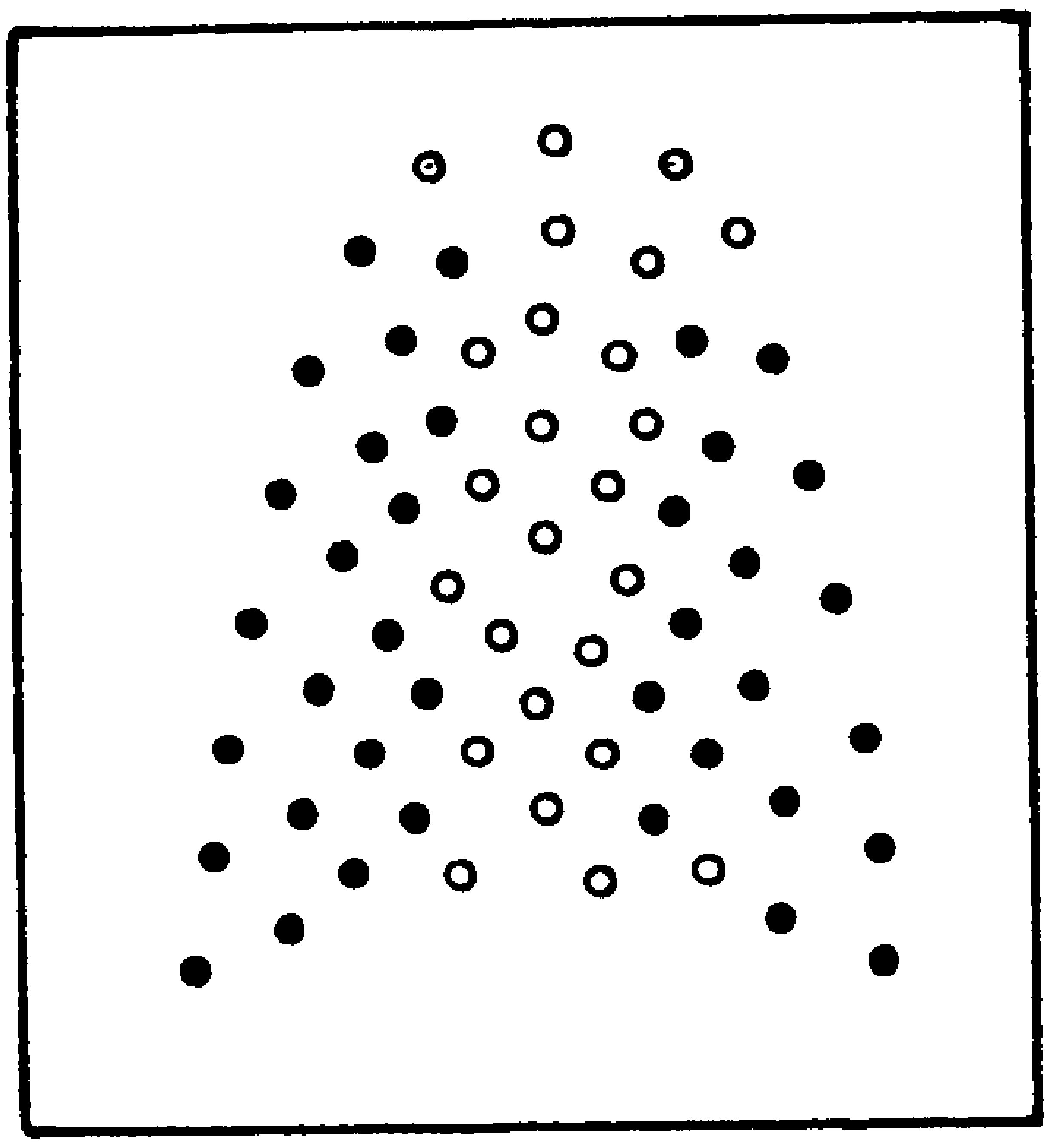
38



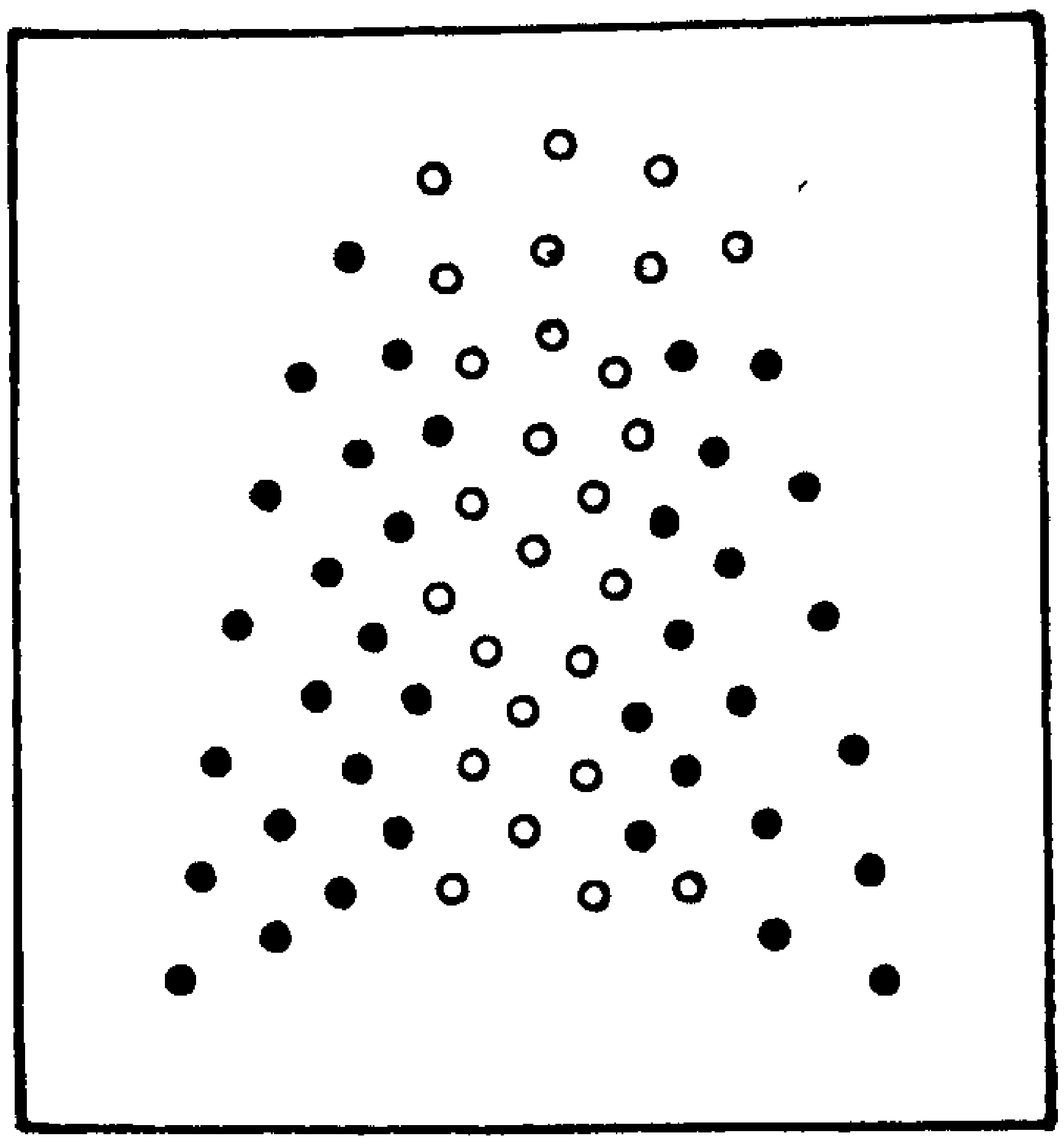
39



40



41



42

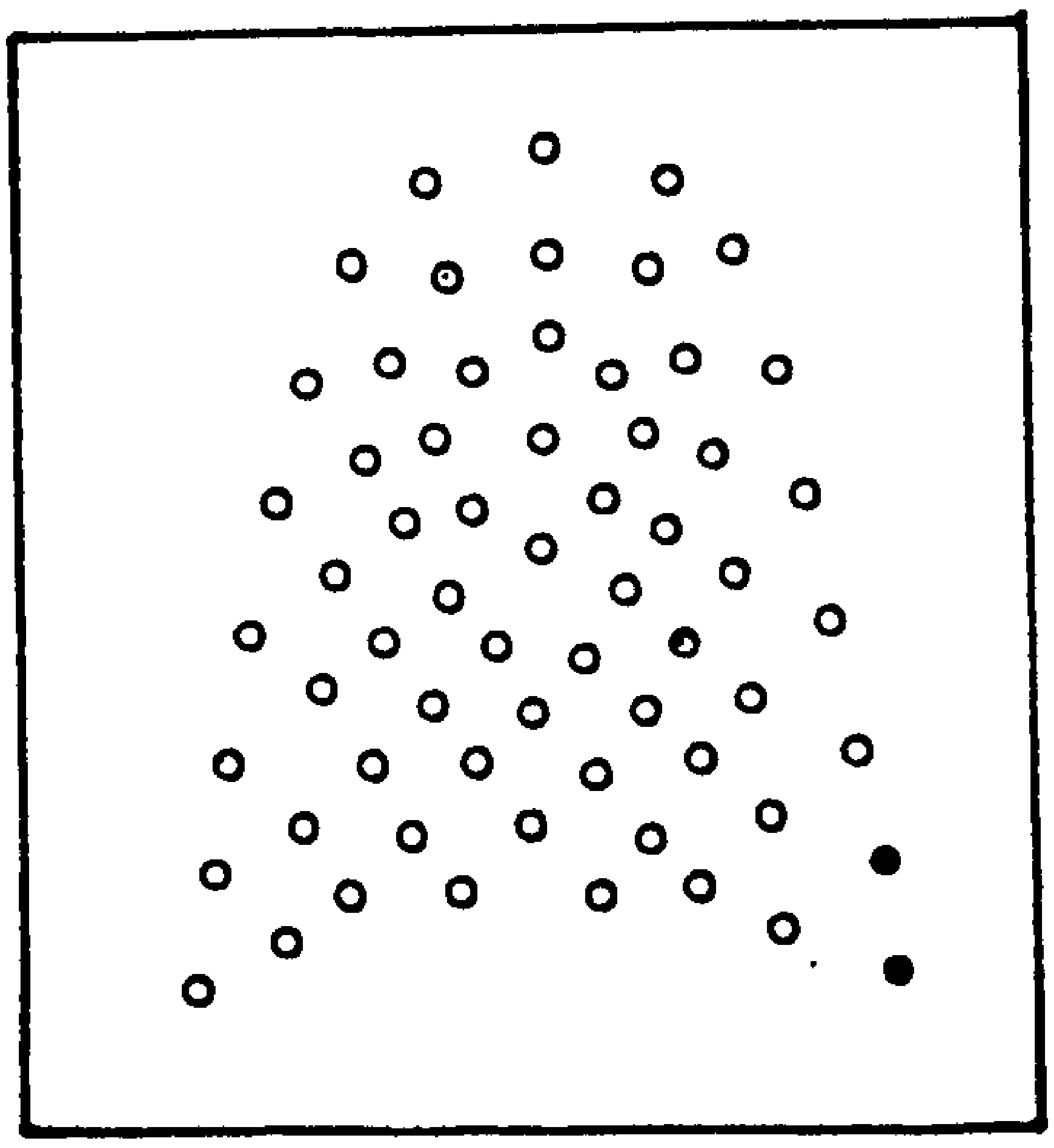


Fig. (5.6) Diagrams showing EPG patterns for the single plosive [d] in the environment of /ξada/.

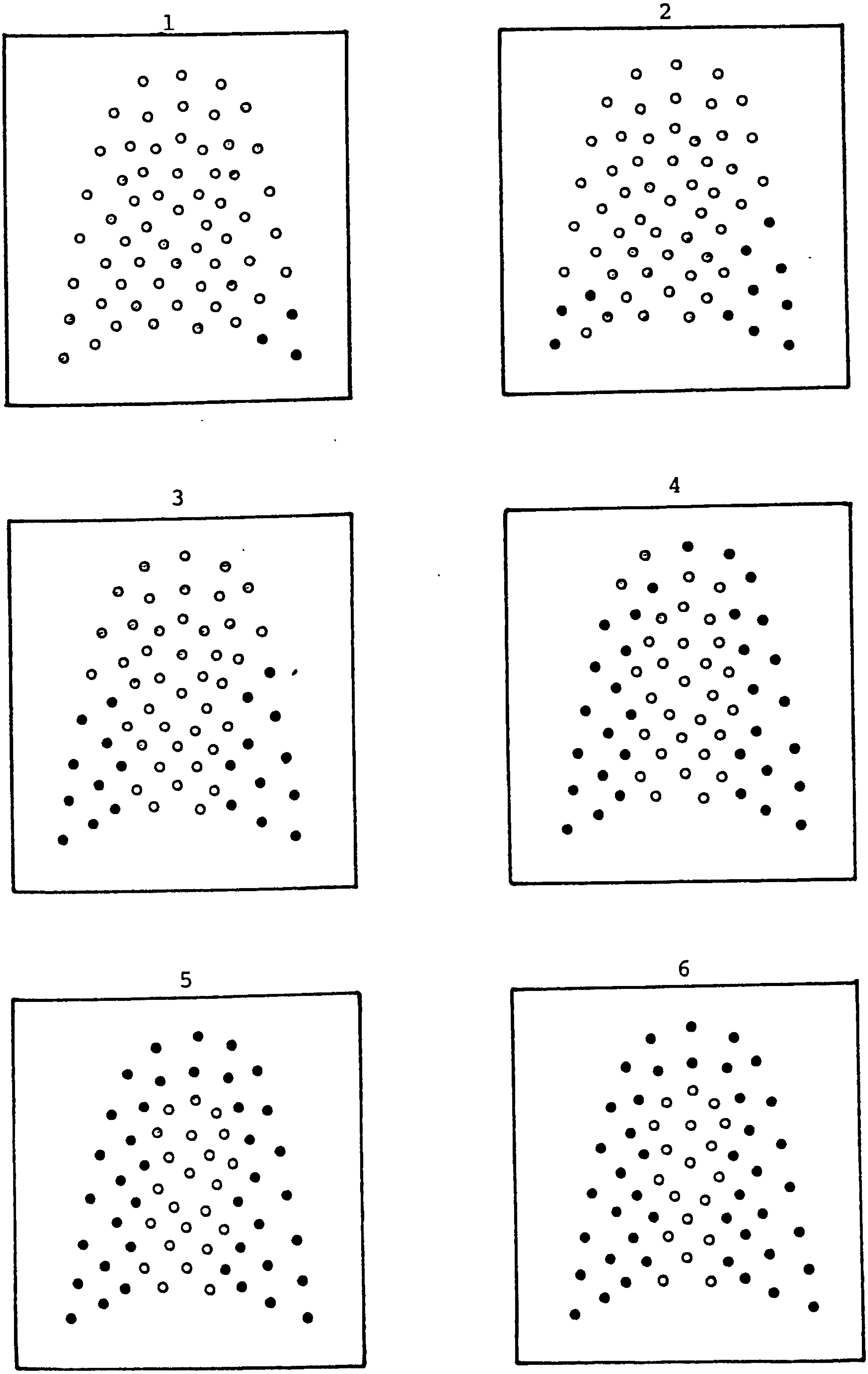


Fig. (5.6) cont'd

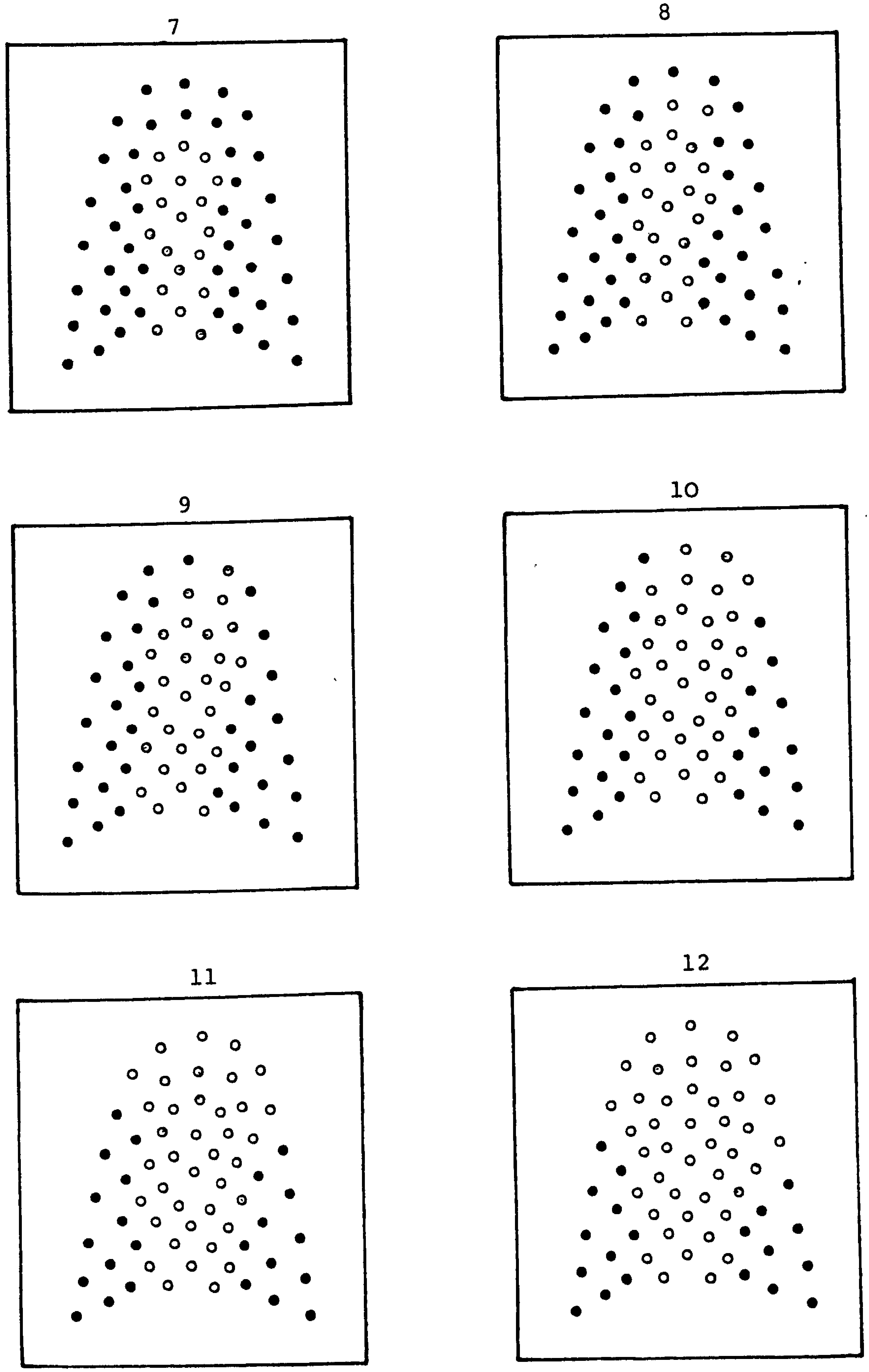


Fig. (5.6) cont'd

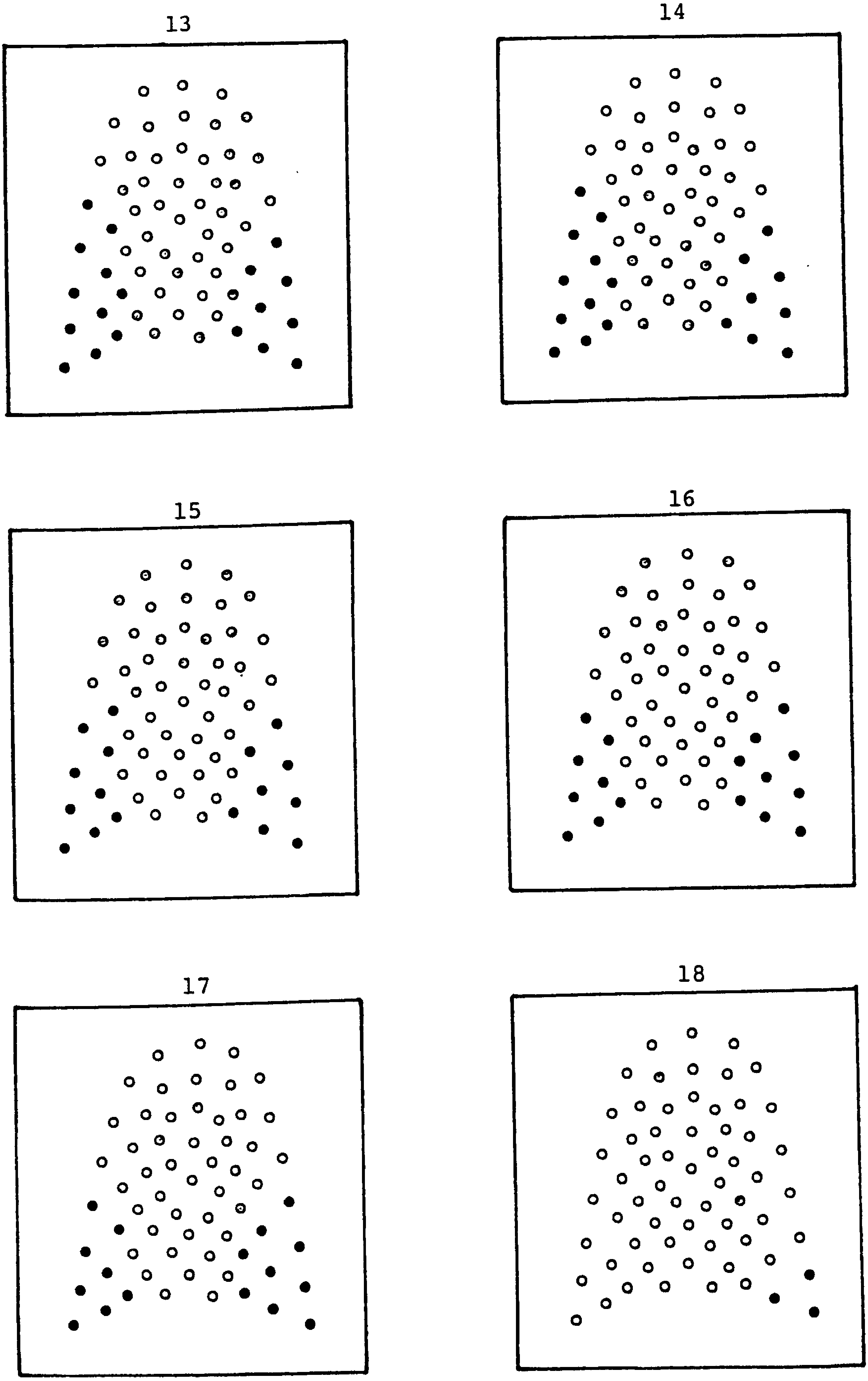


Fig. (5.7) Diagrams showing EPG patterns for the geminate plosive [dd] in the environment of /ʕadda/.

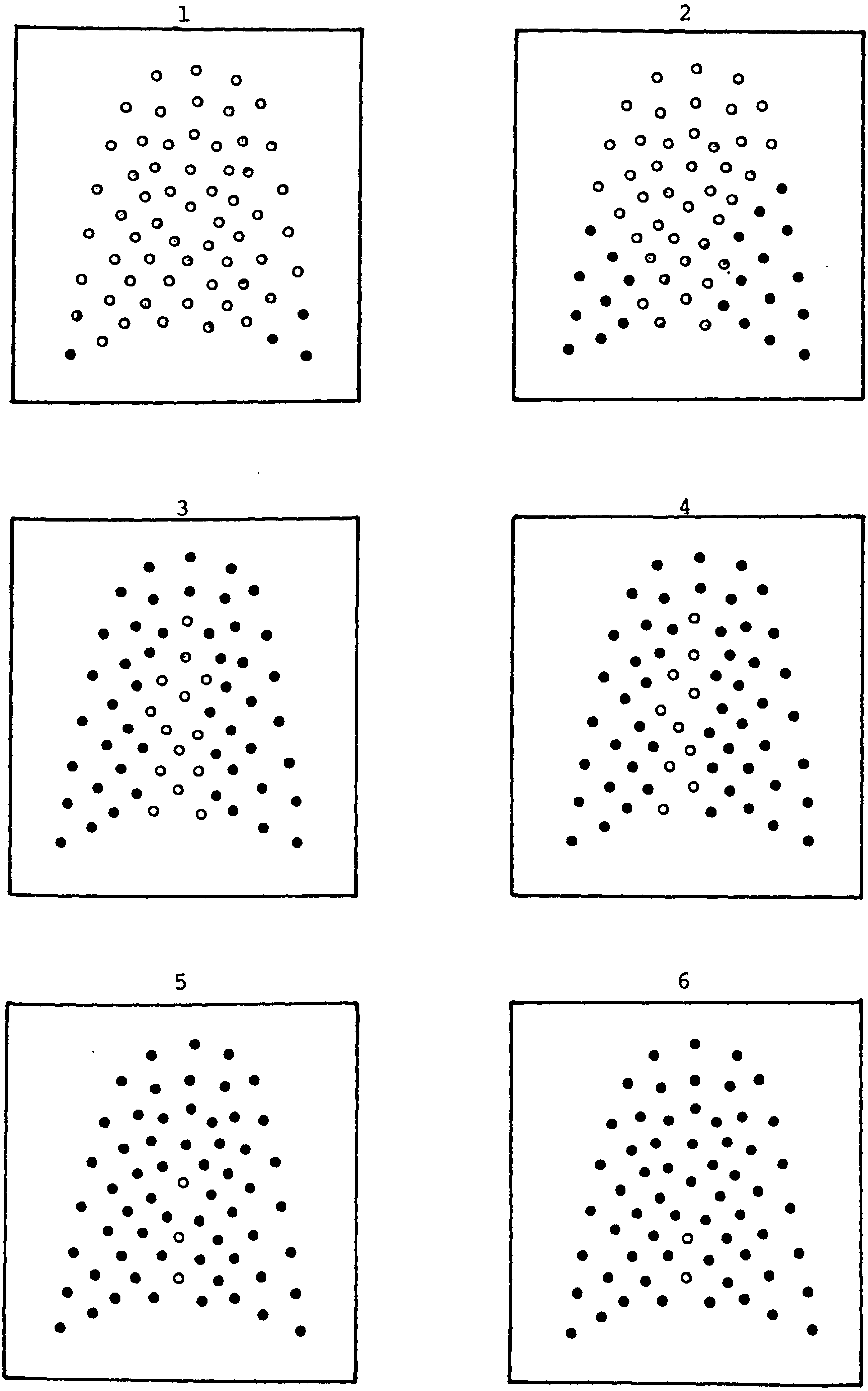
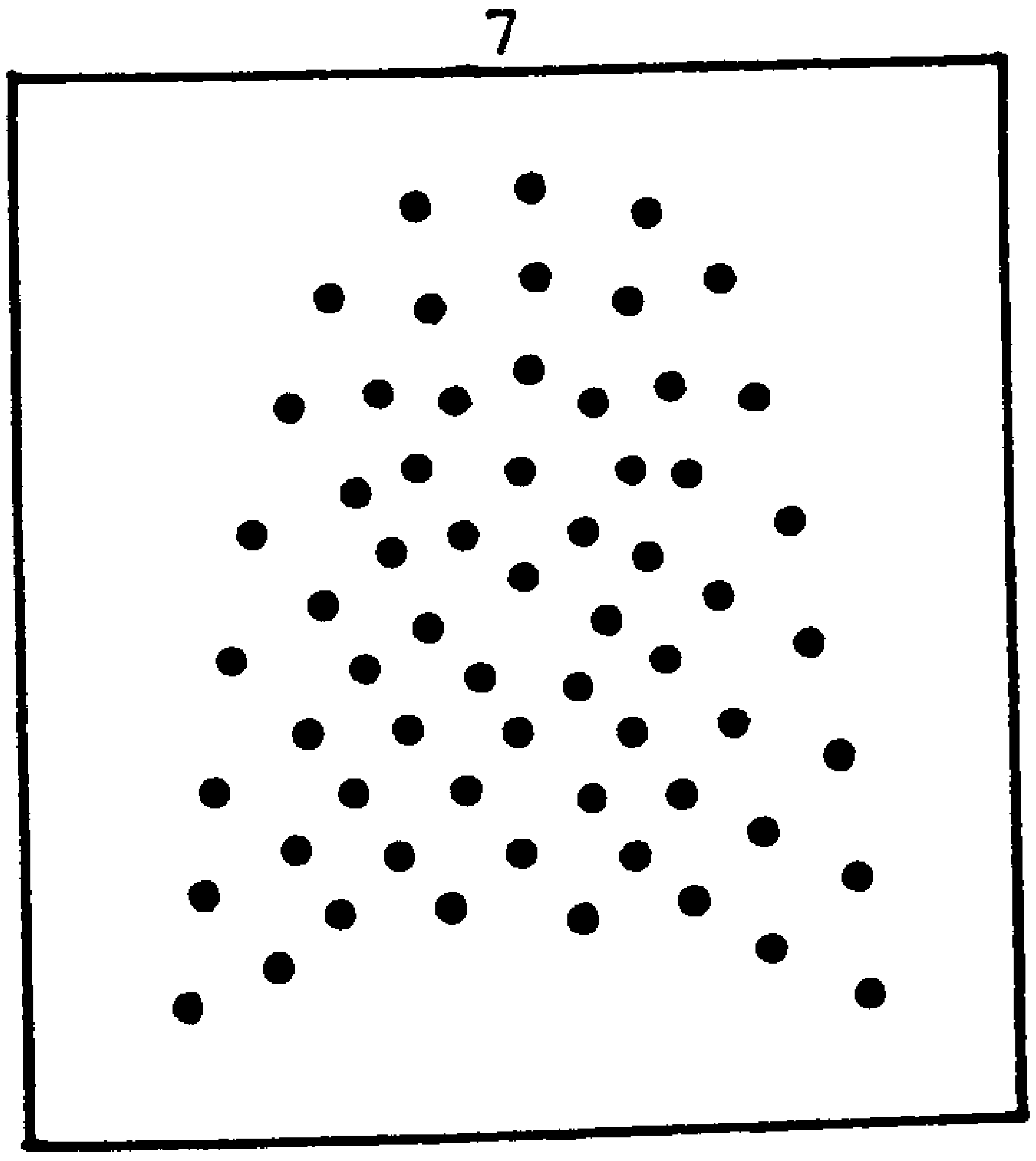
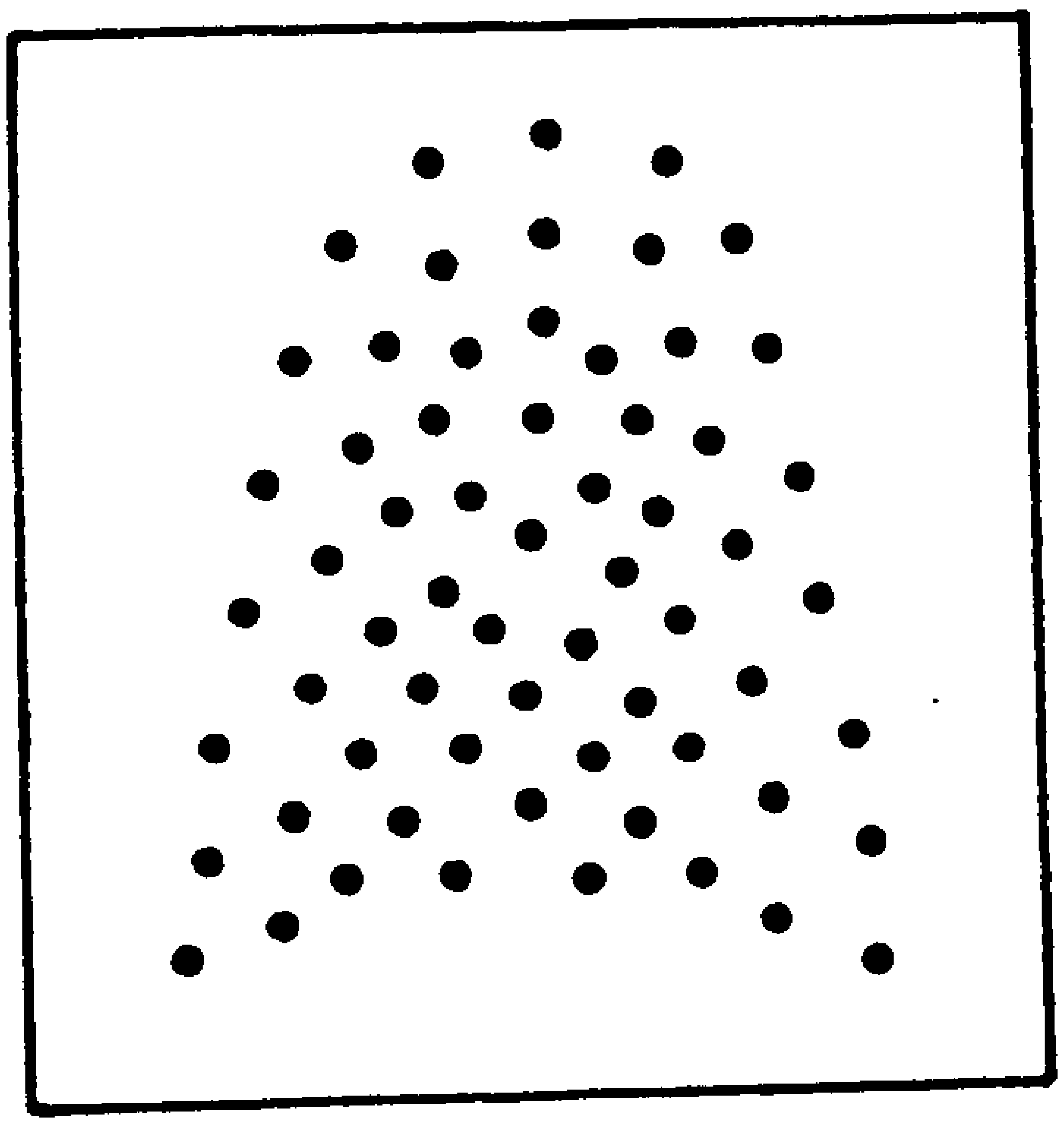


Fig. (5.7) cont'd

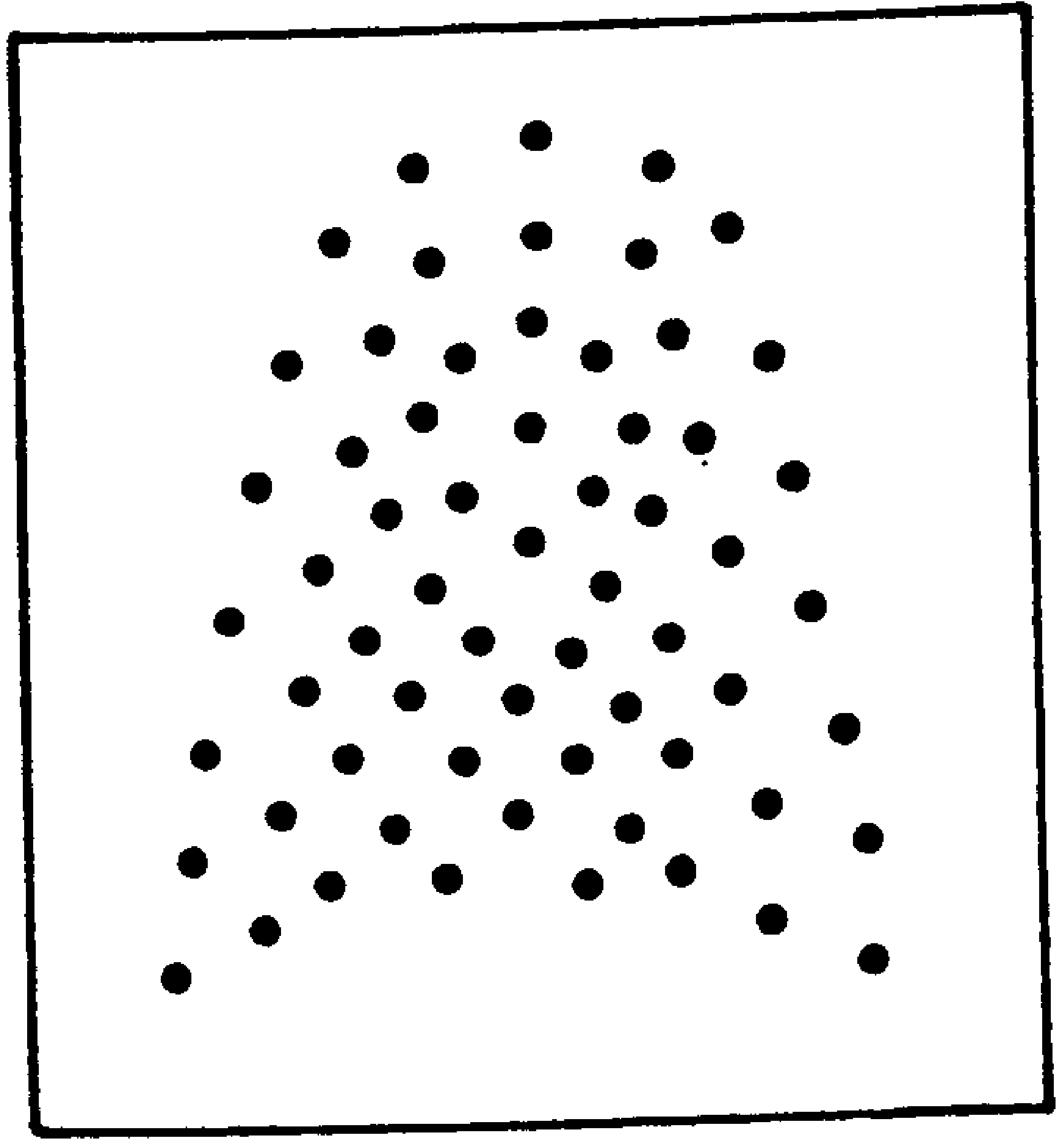
7



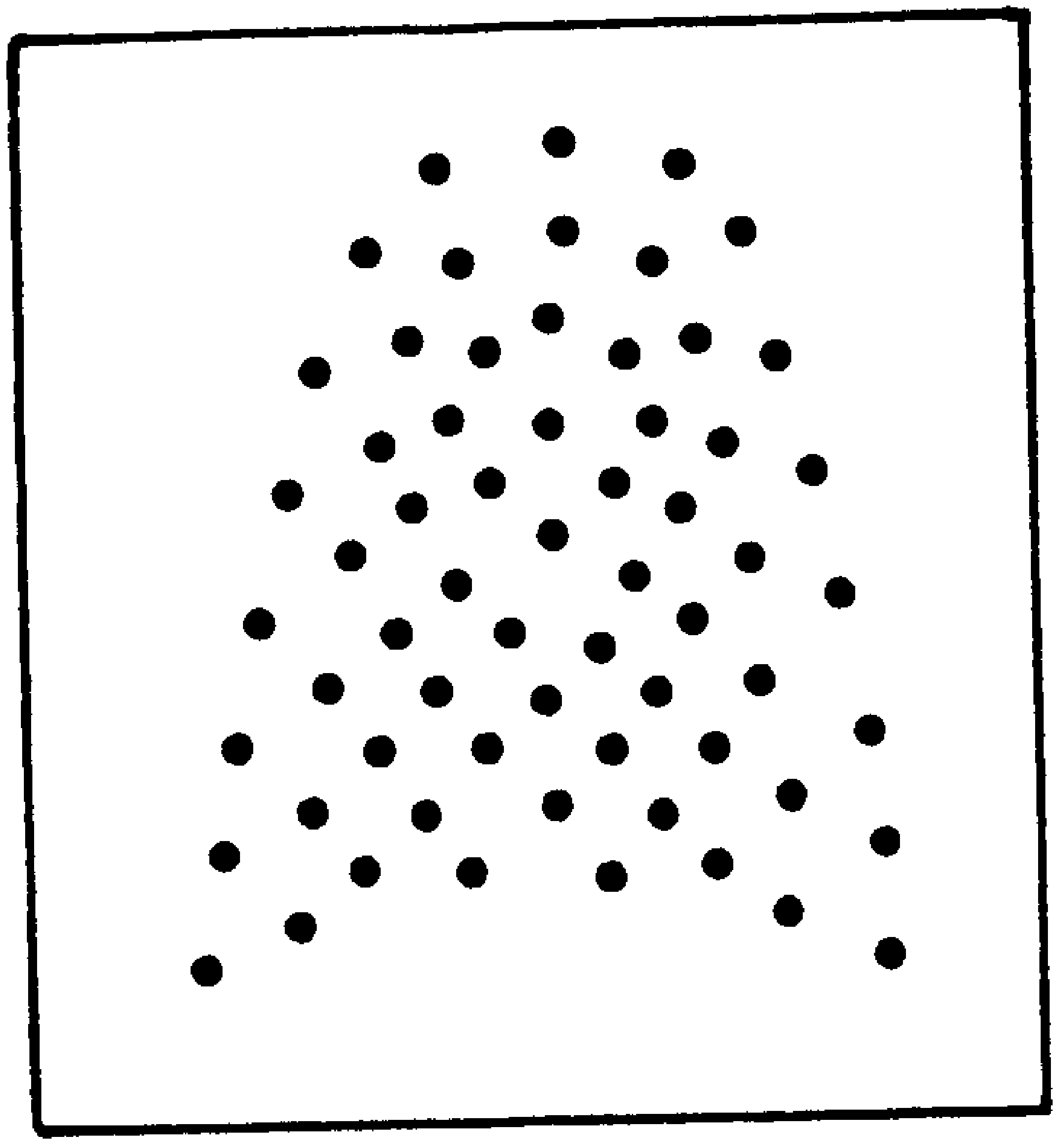
8



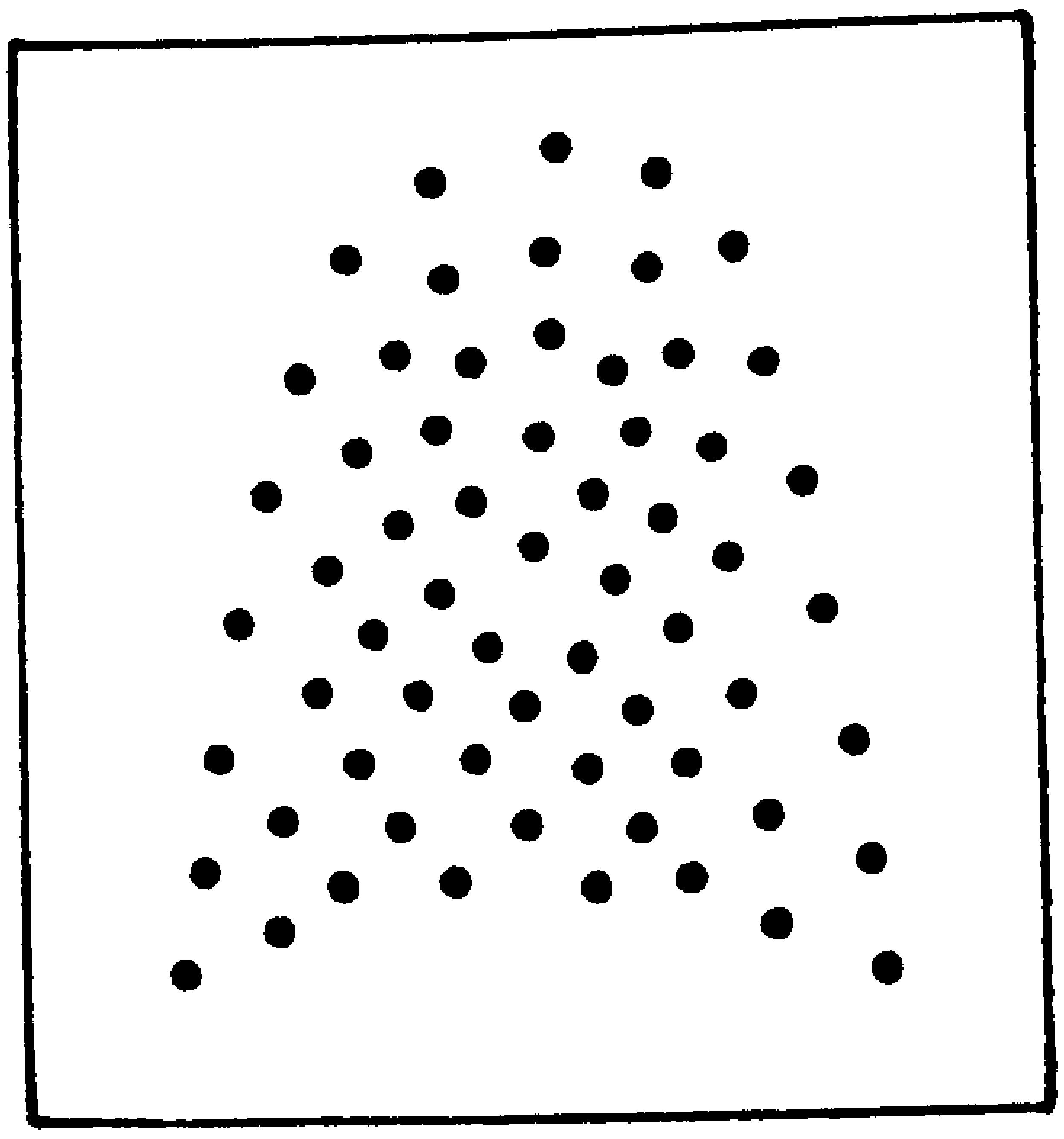
9



10



11



12

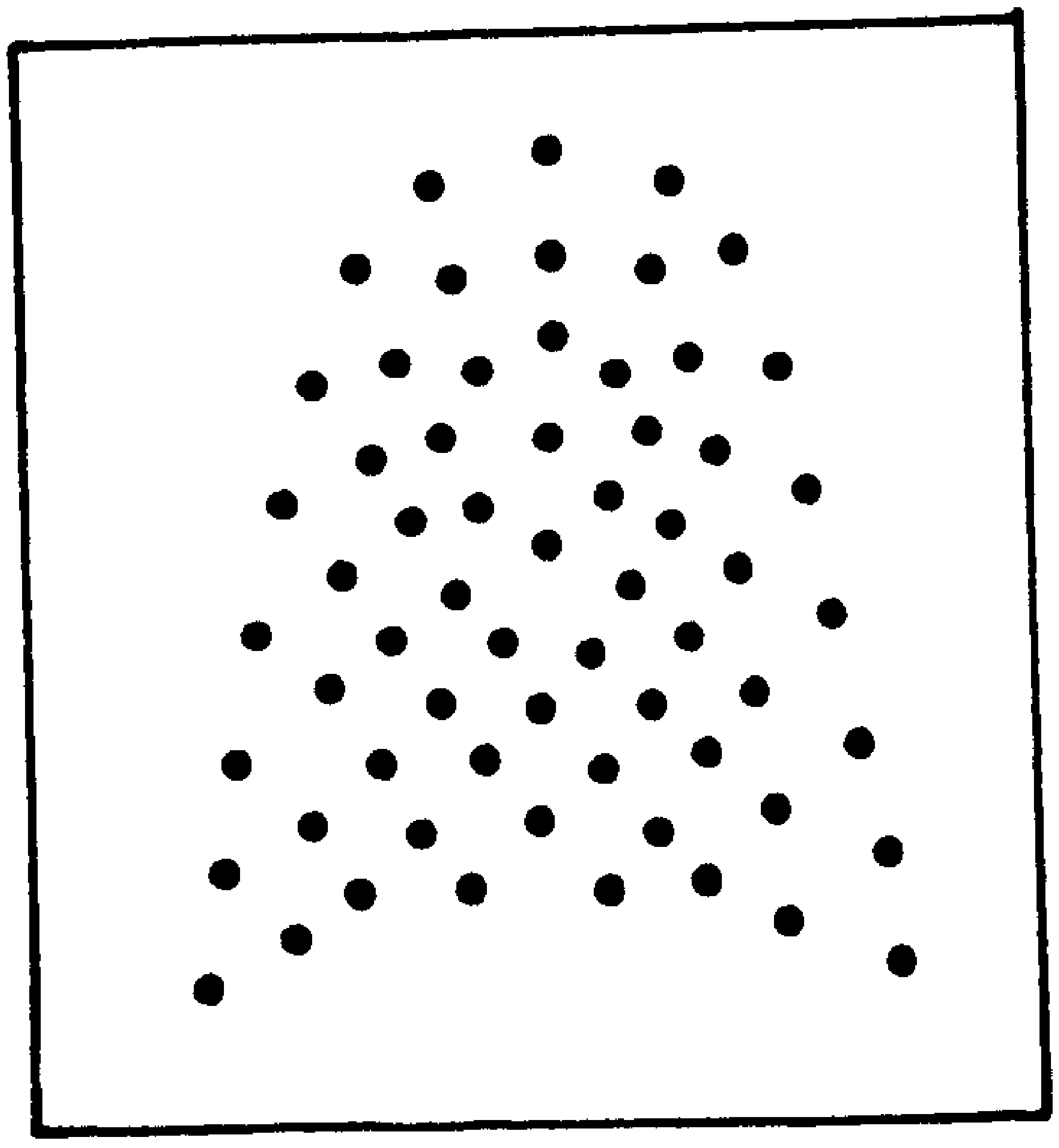
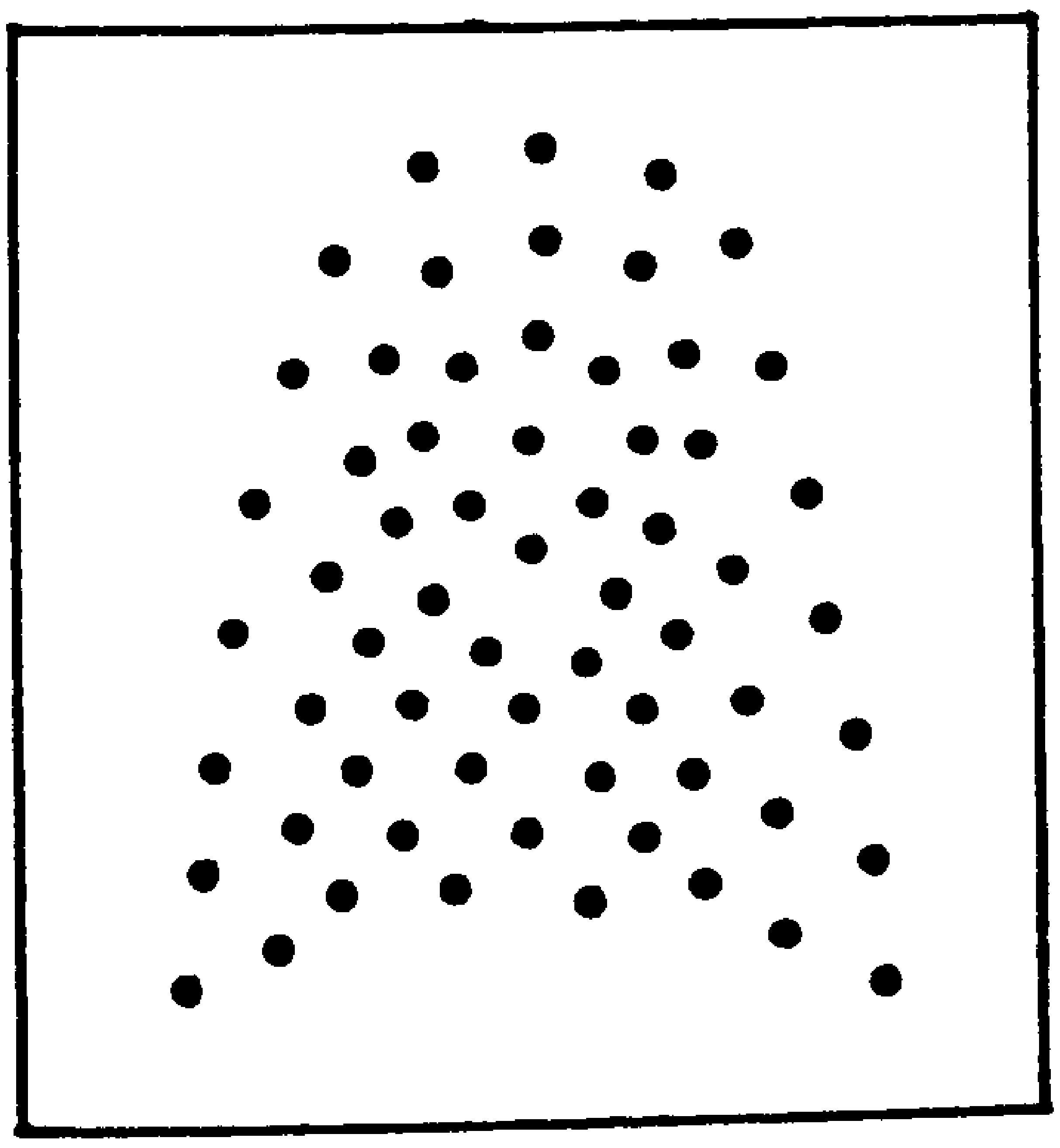
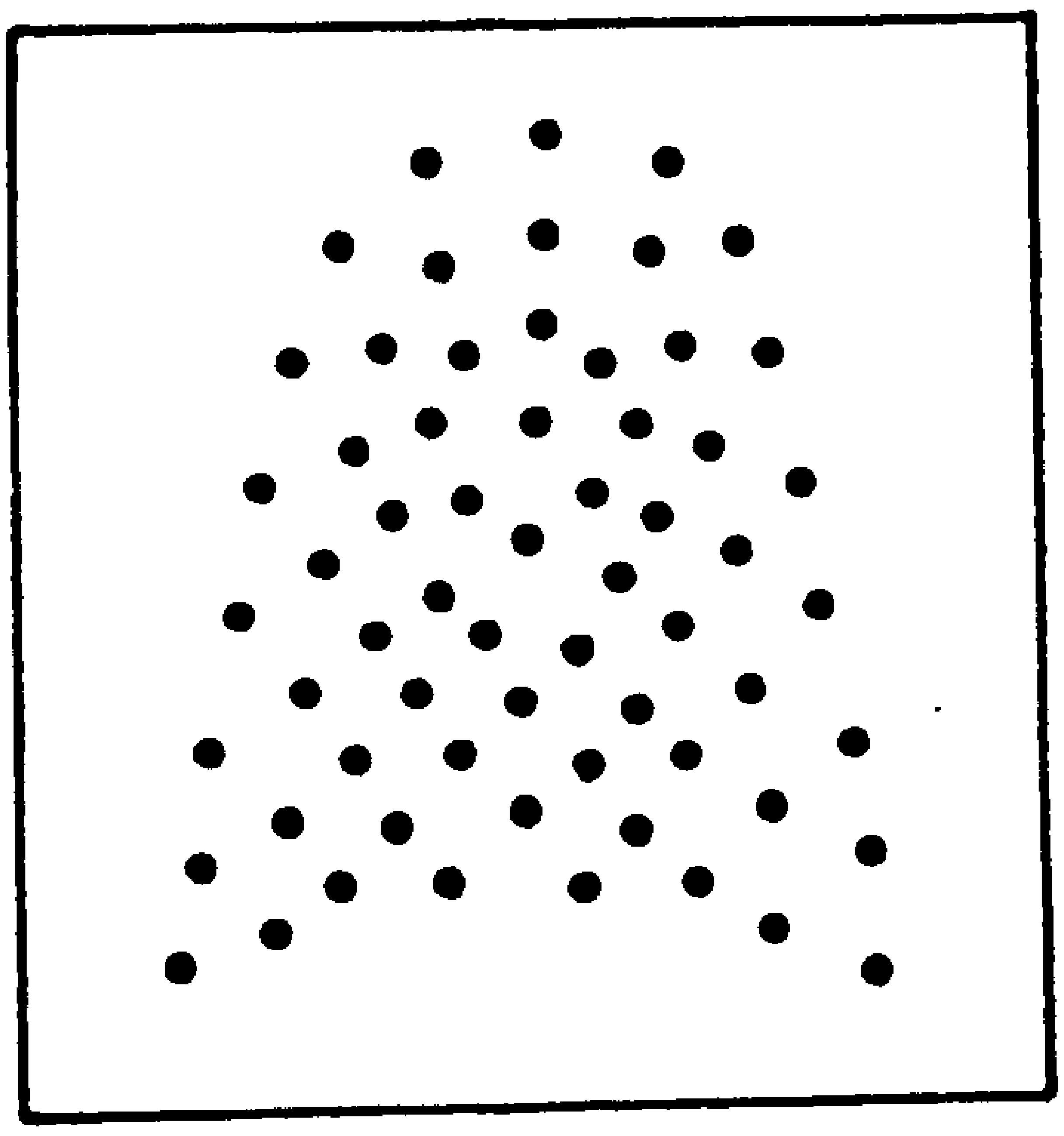


Fig. (5.7) cont'd

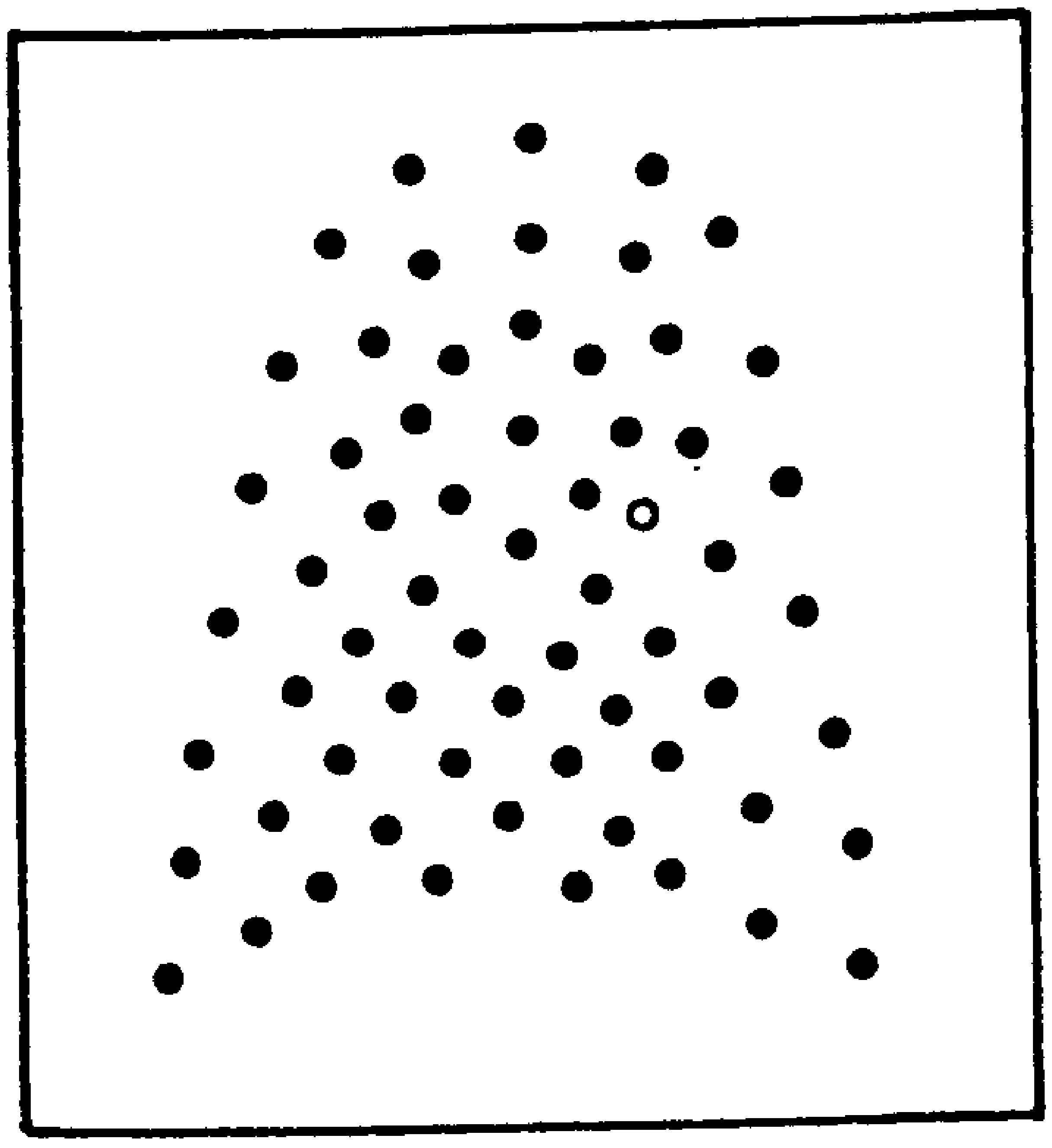
13



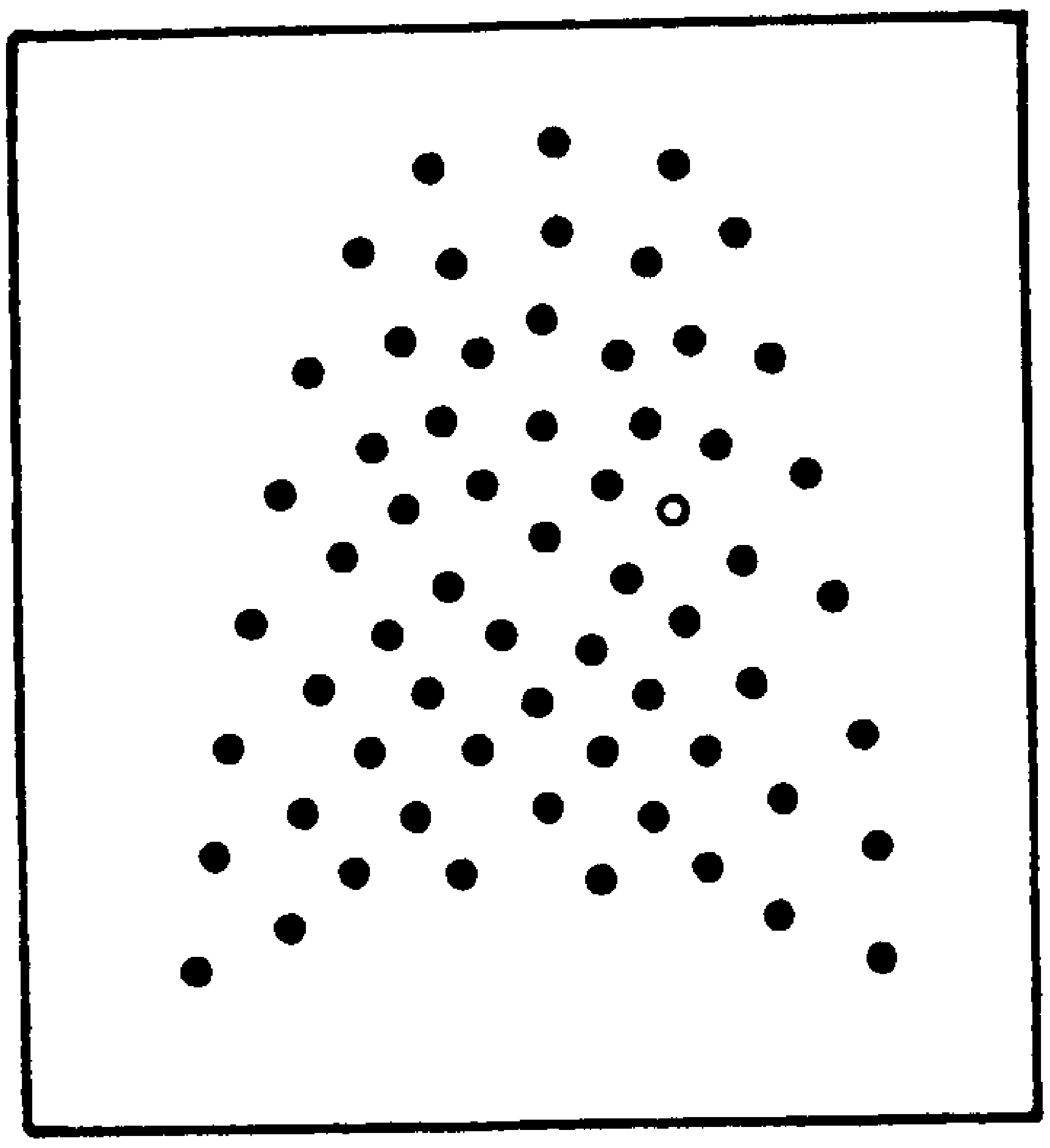
14



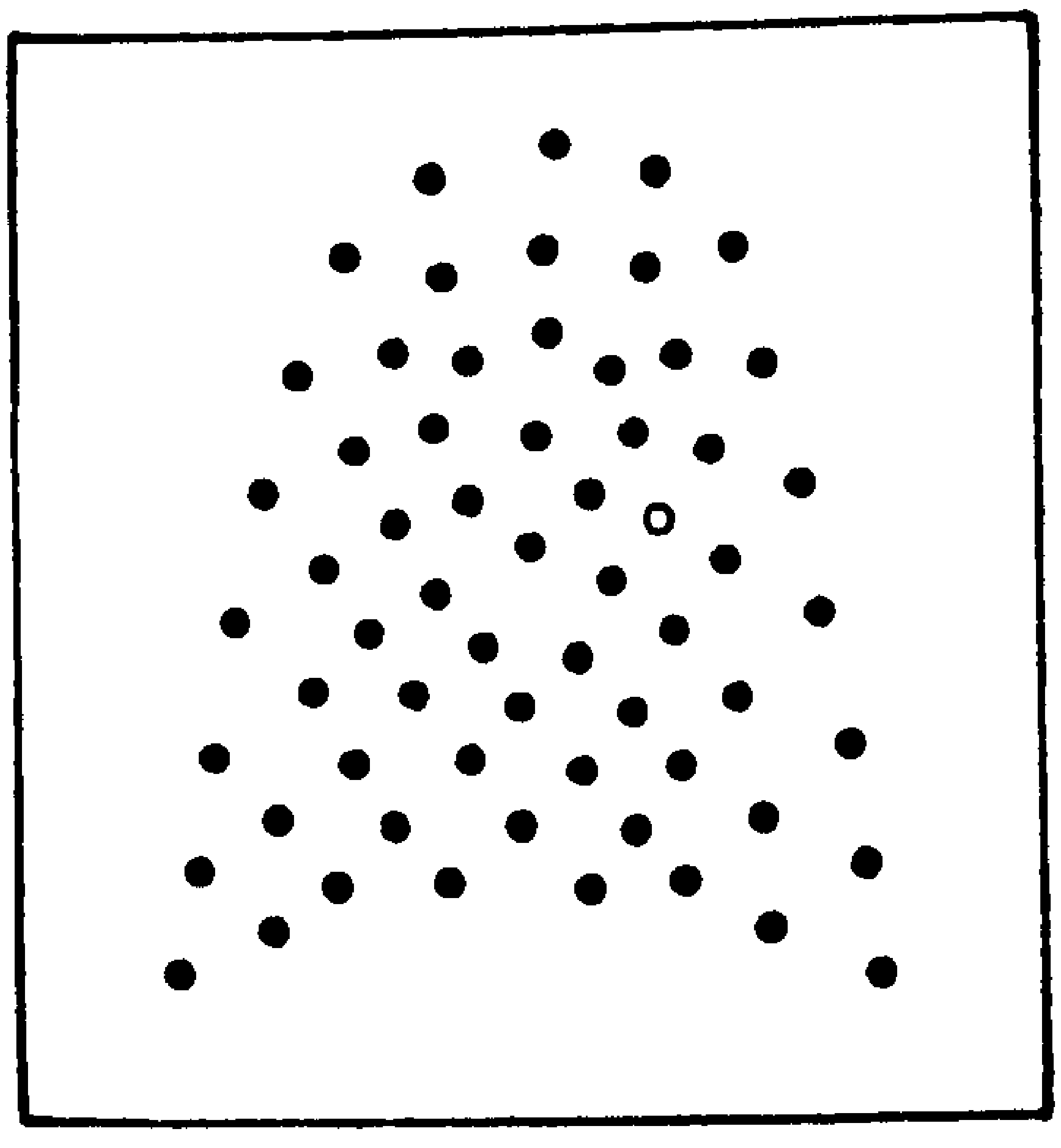
15



16



17



18

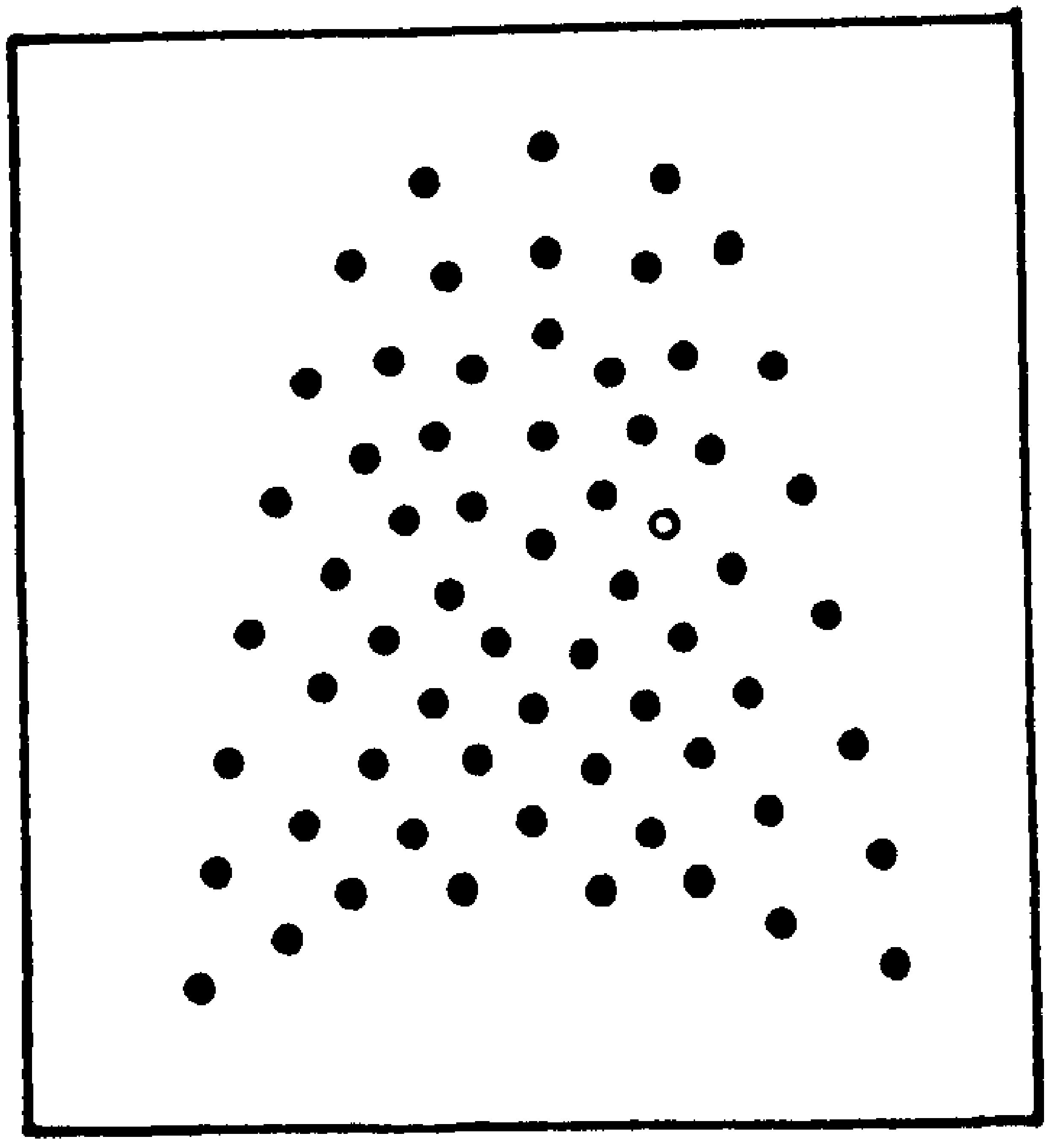
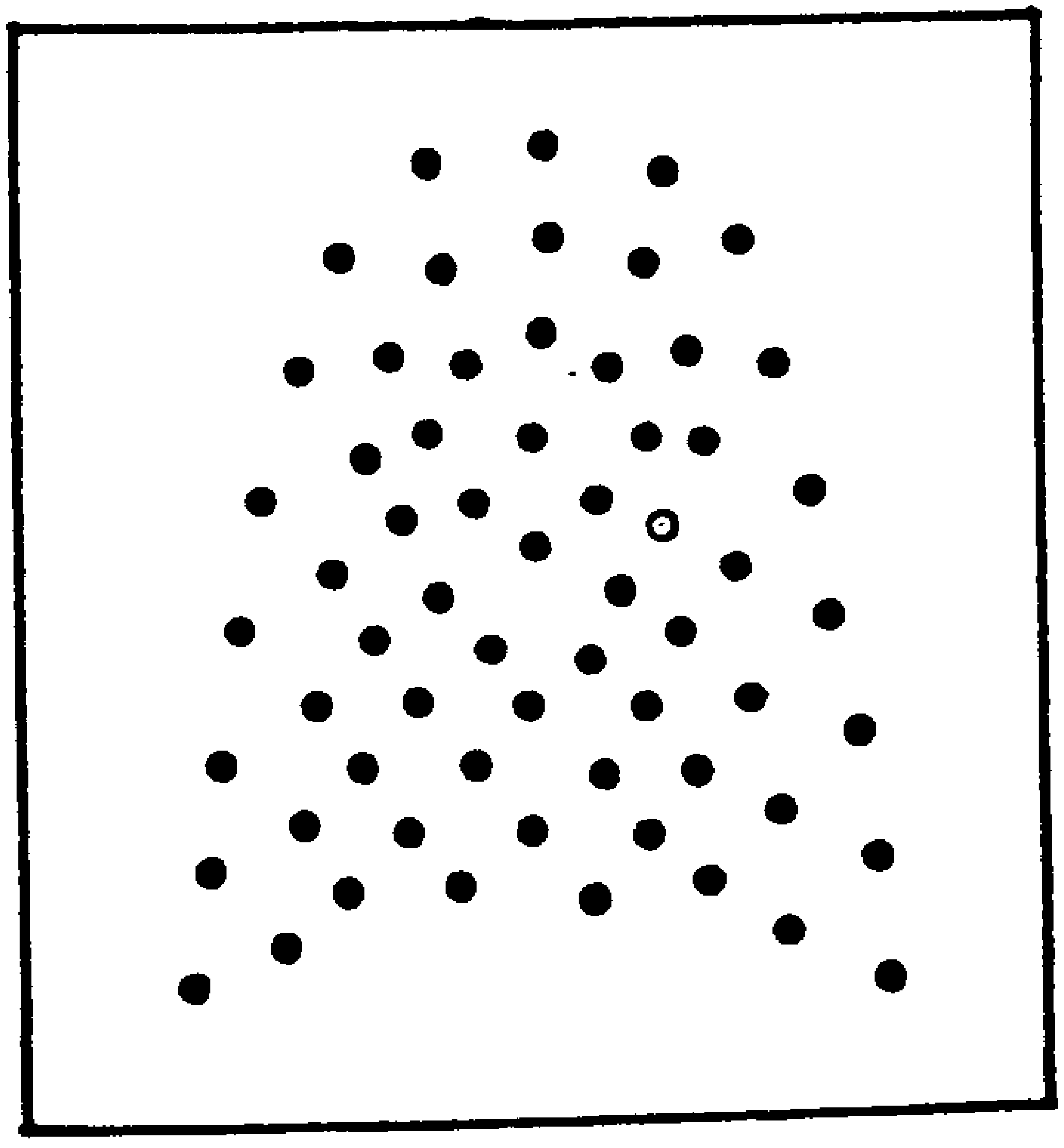
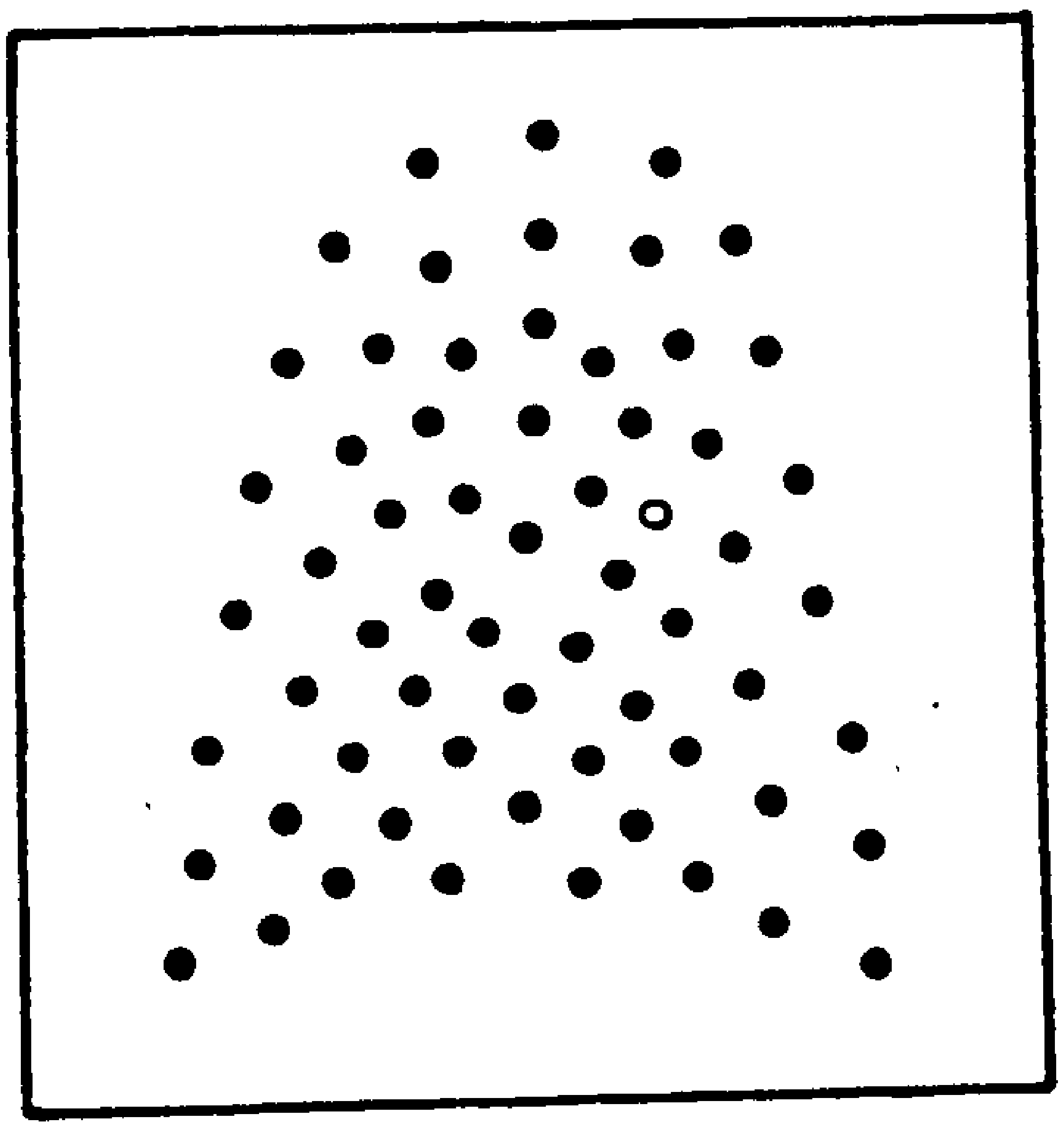


Fig. (5.7) cont'd

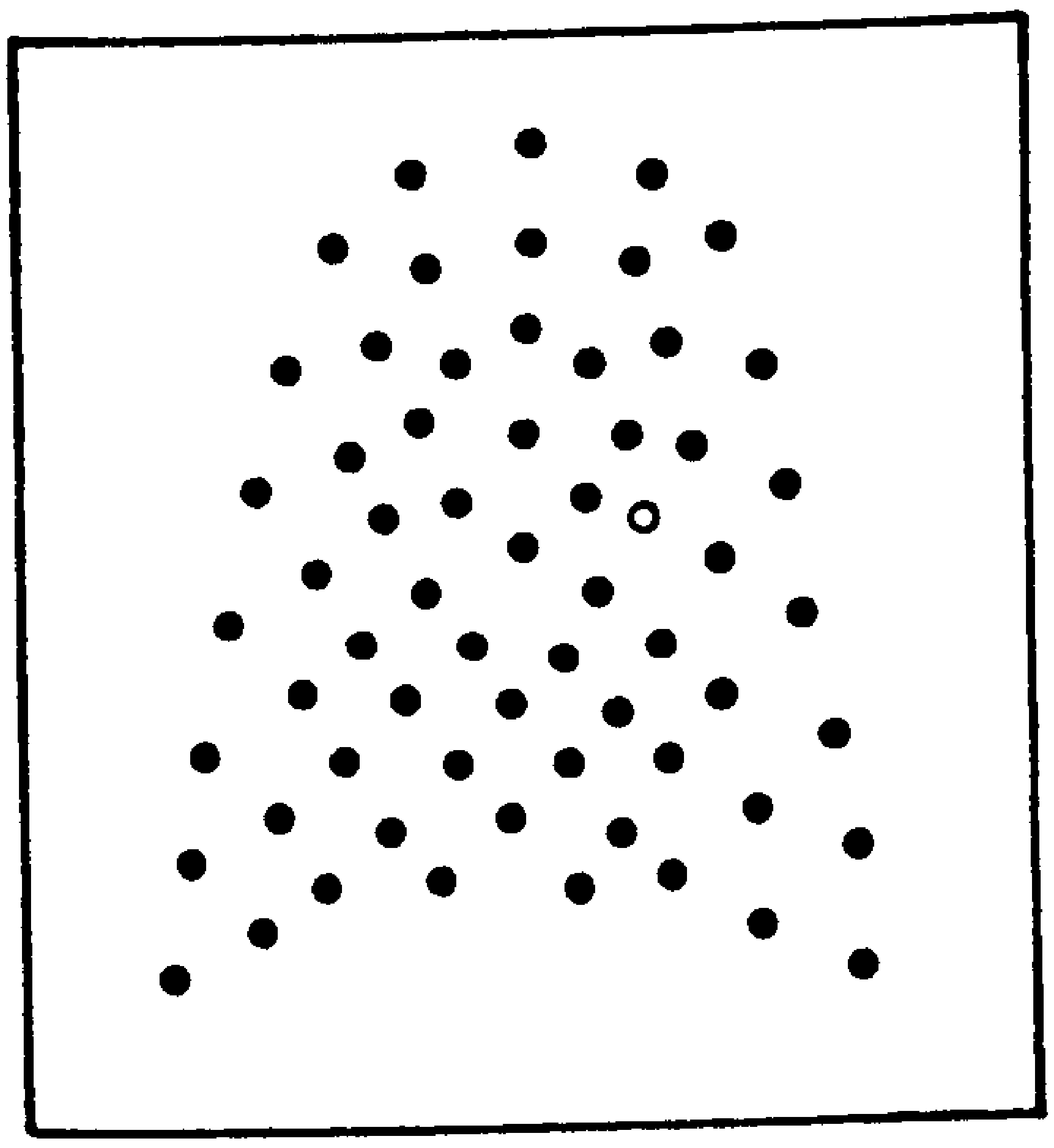
19



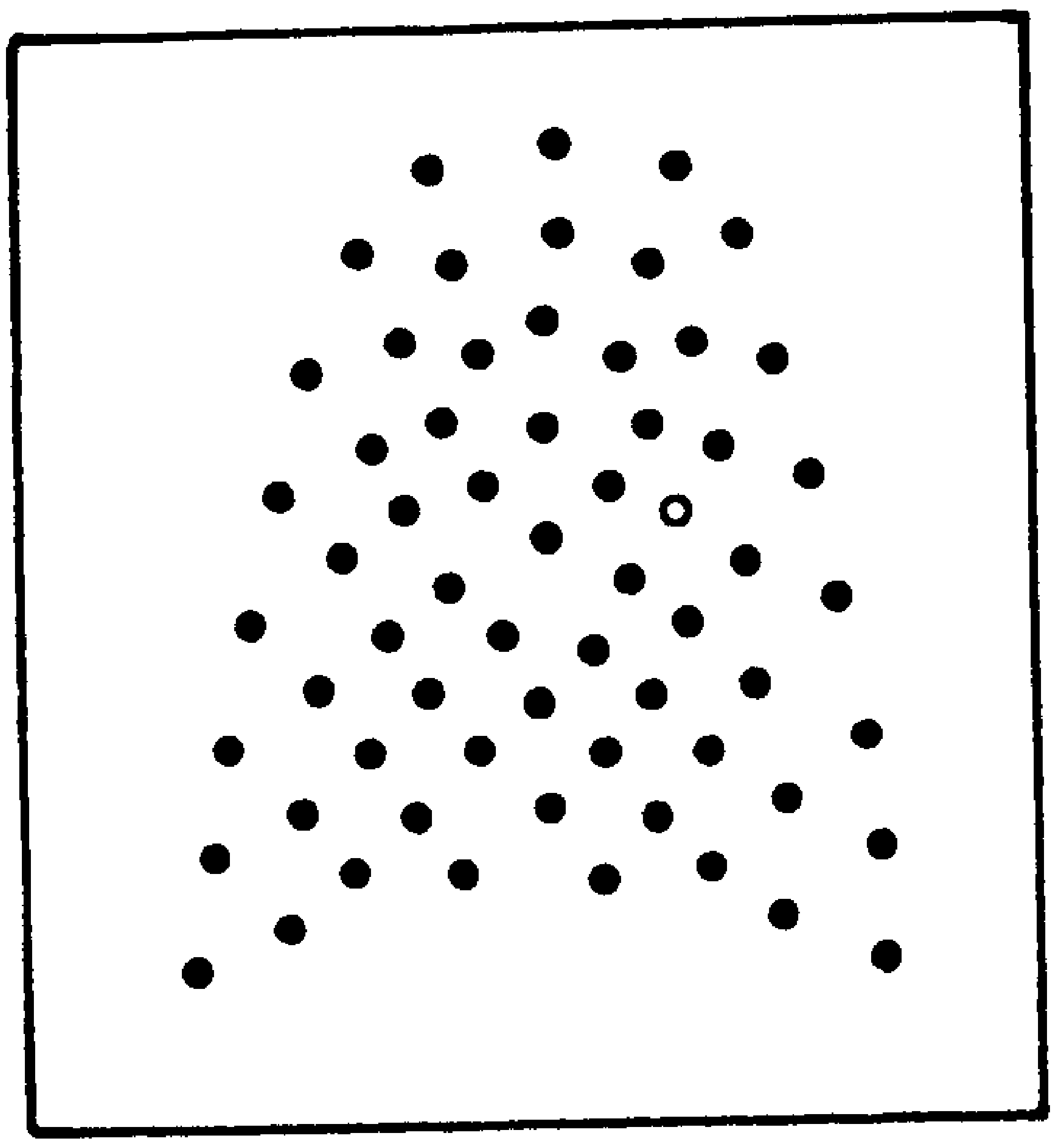
20



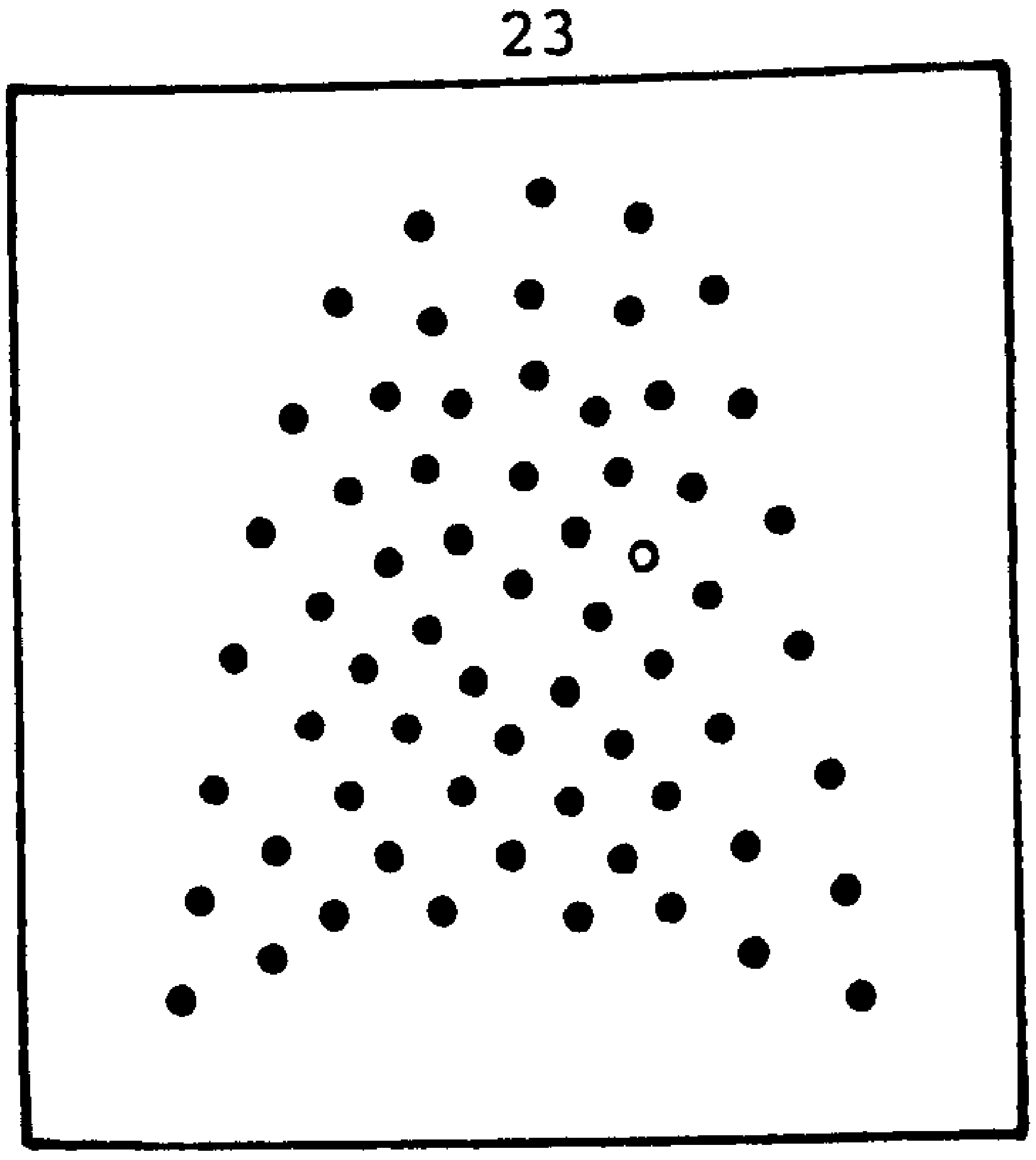
21



22



23



24

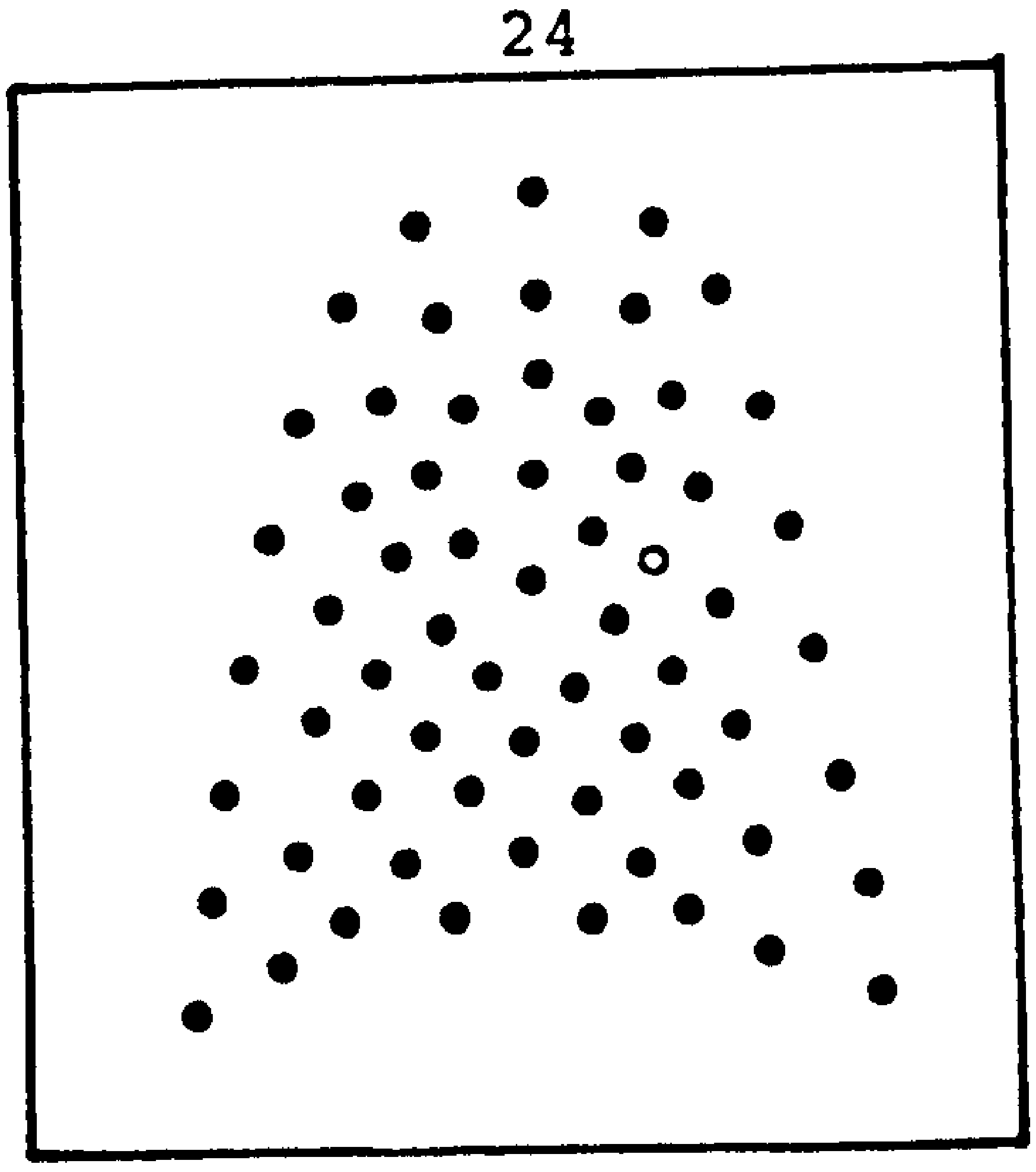
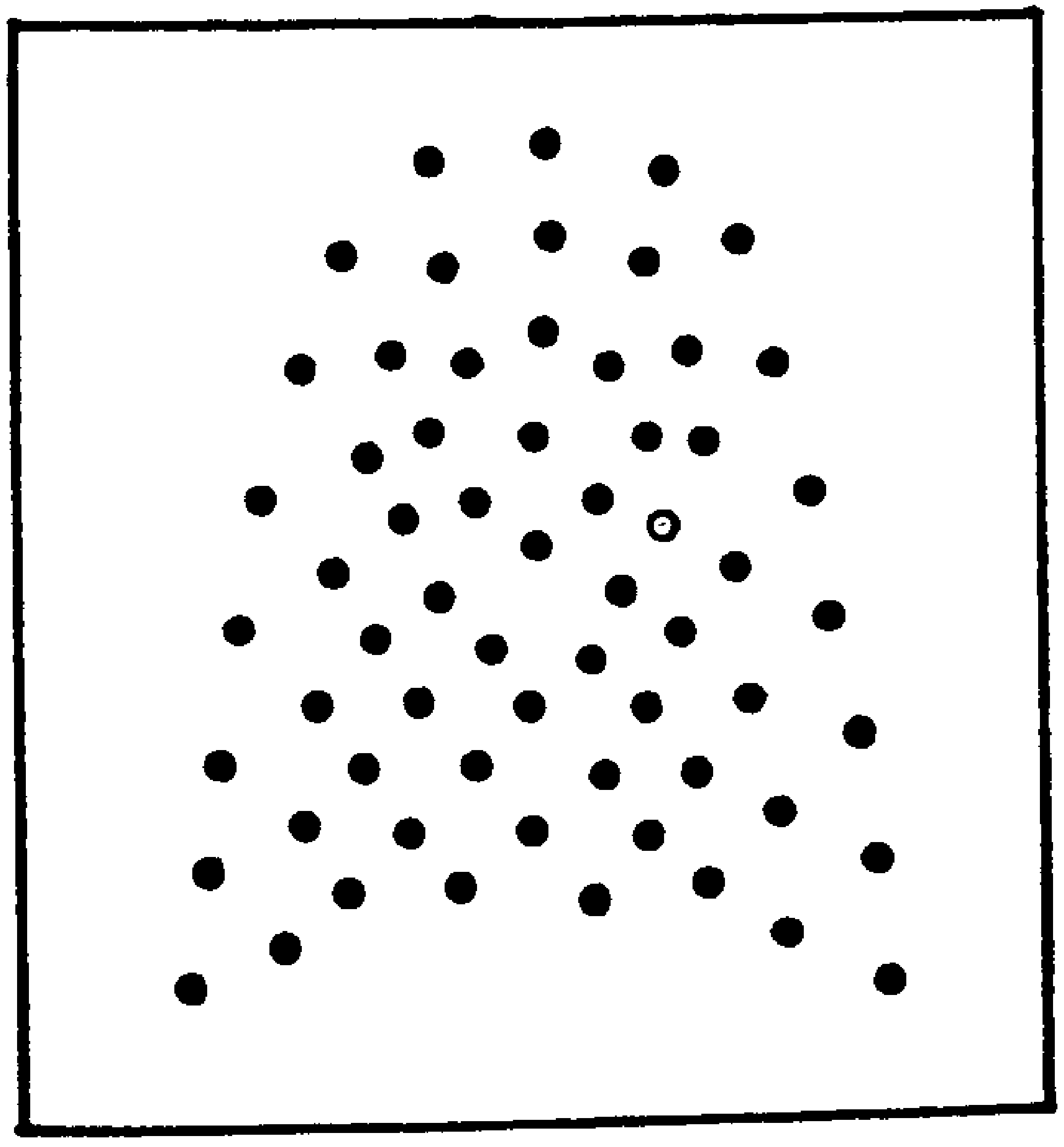
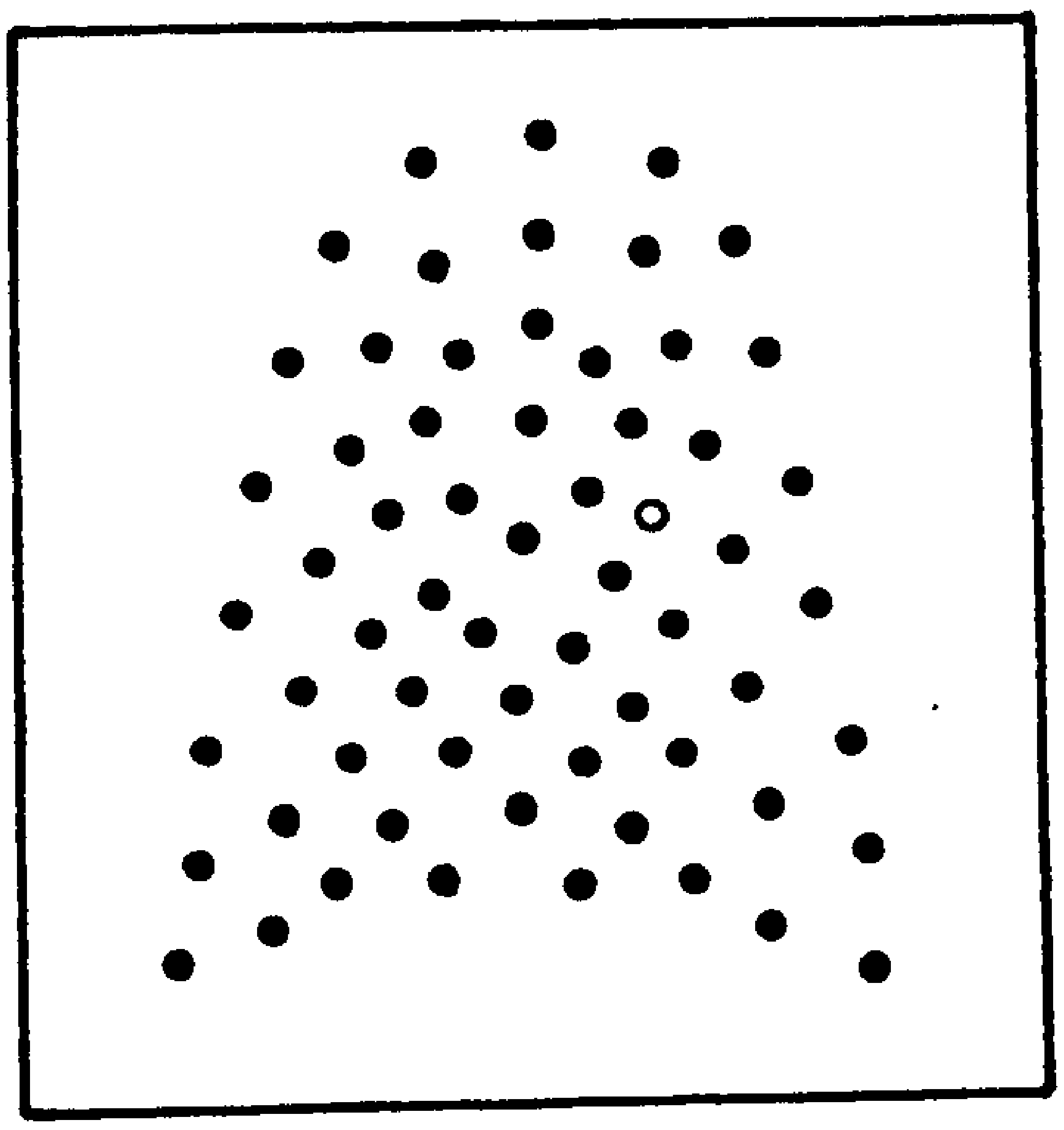


Fig. (5.7) cont'd

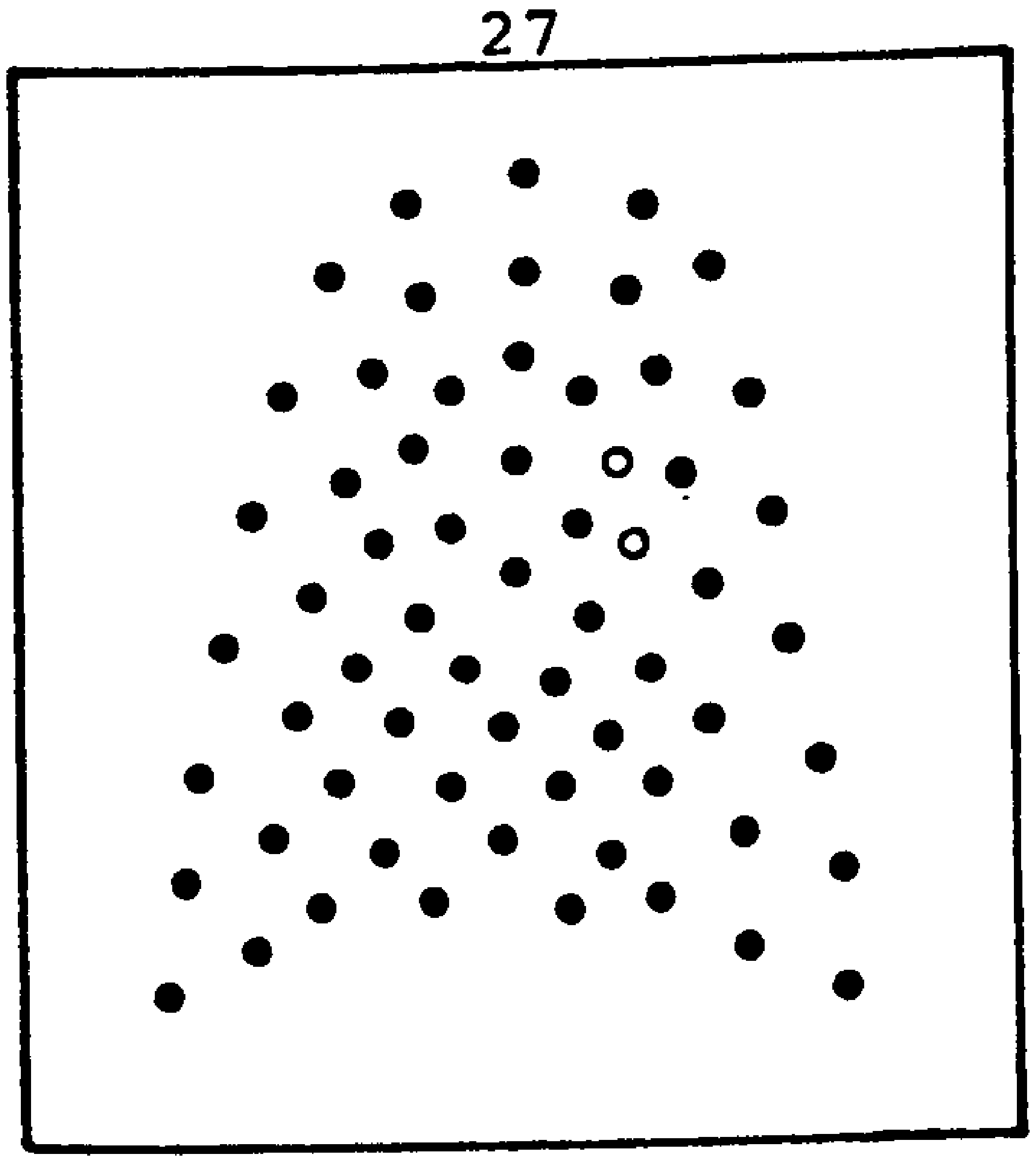
25



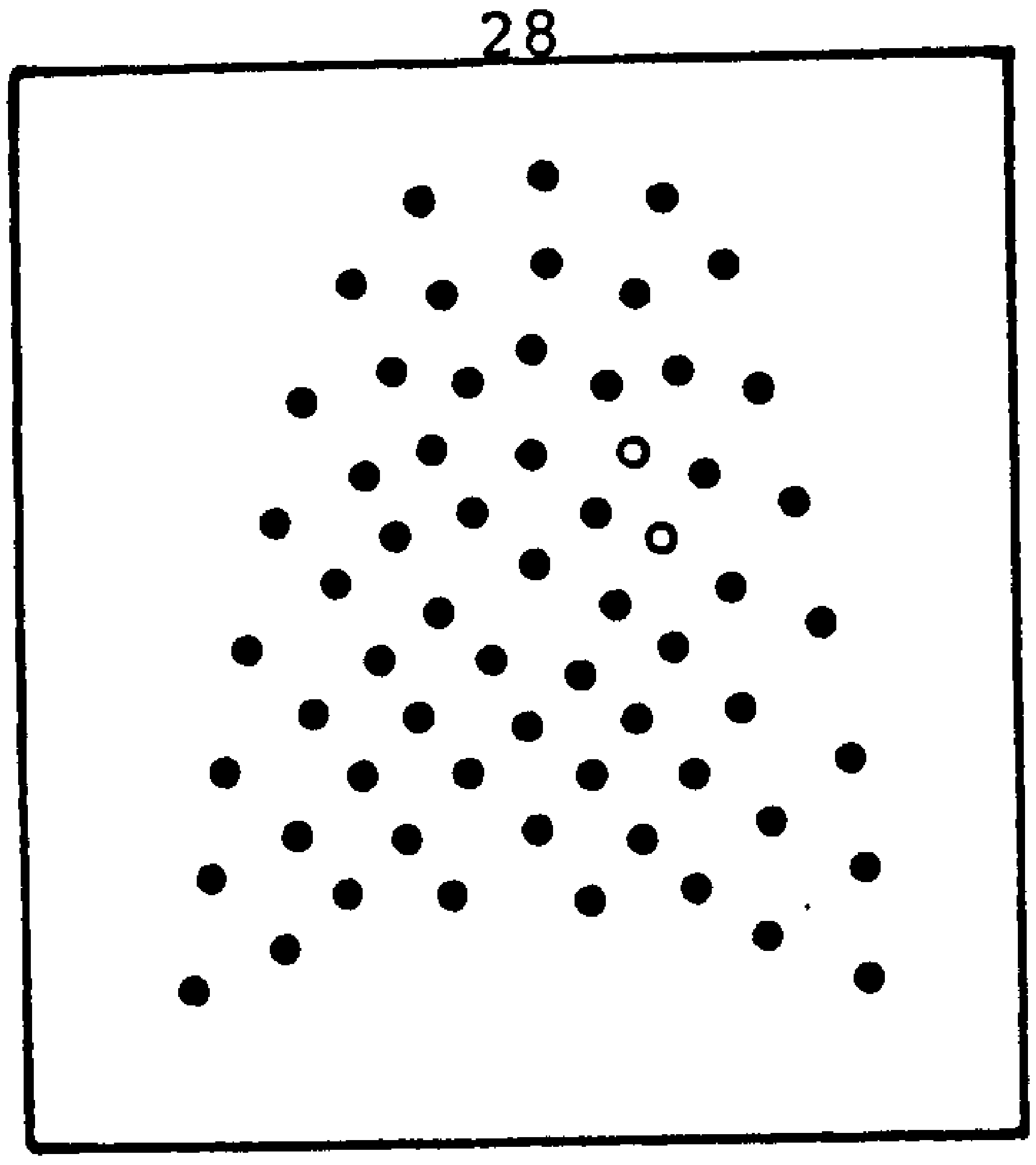
26



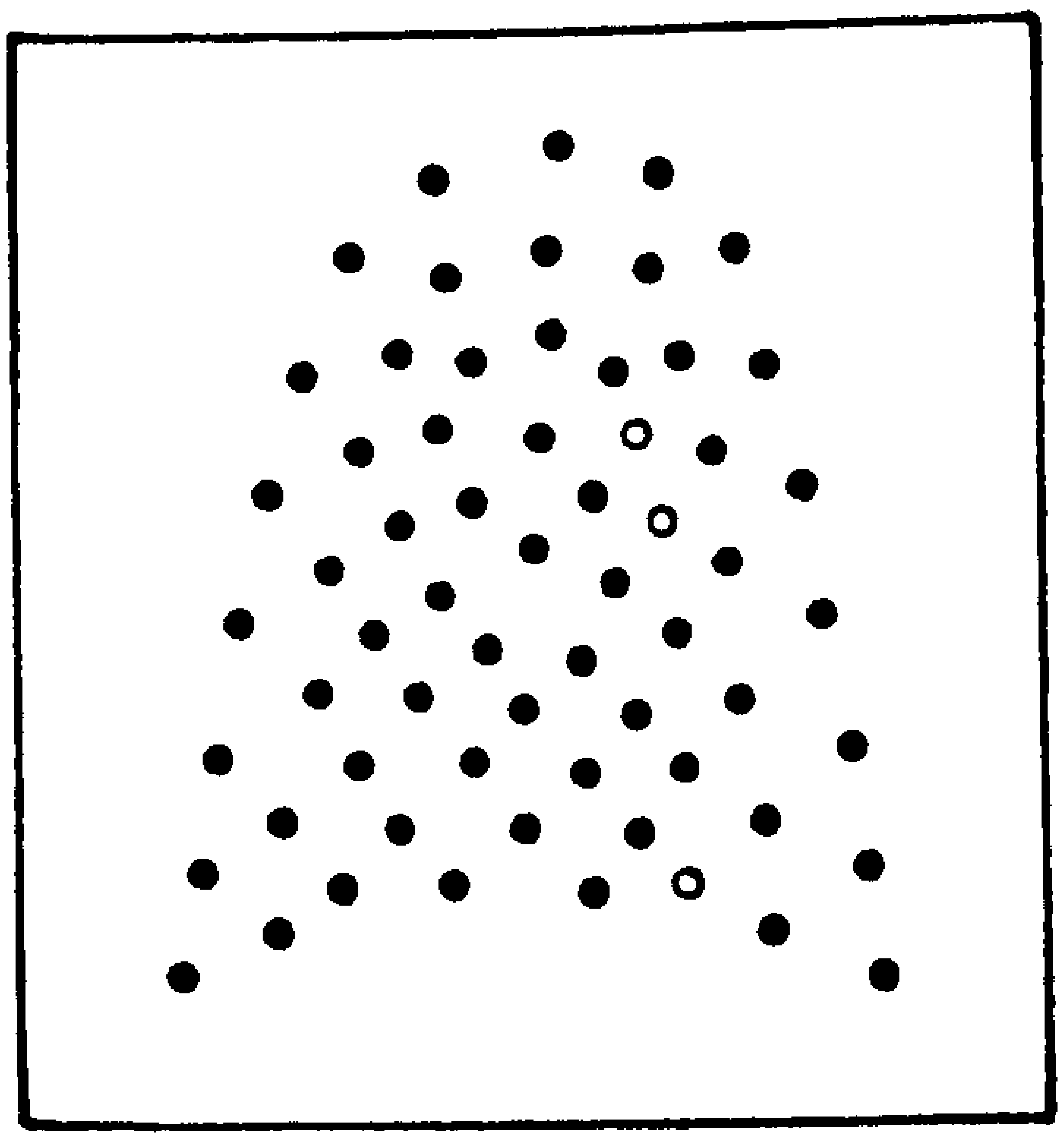
27



28



29



30

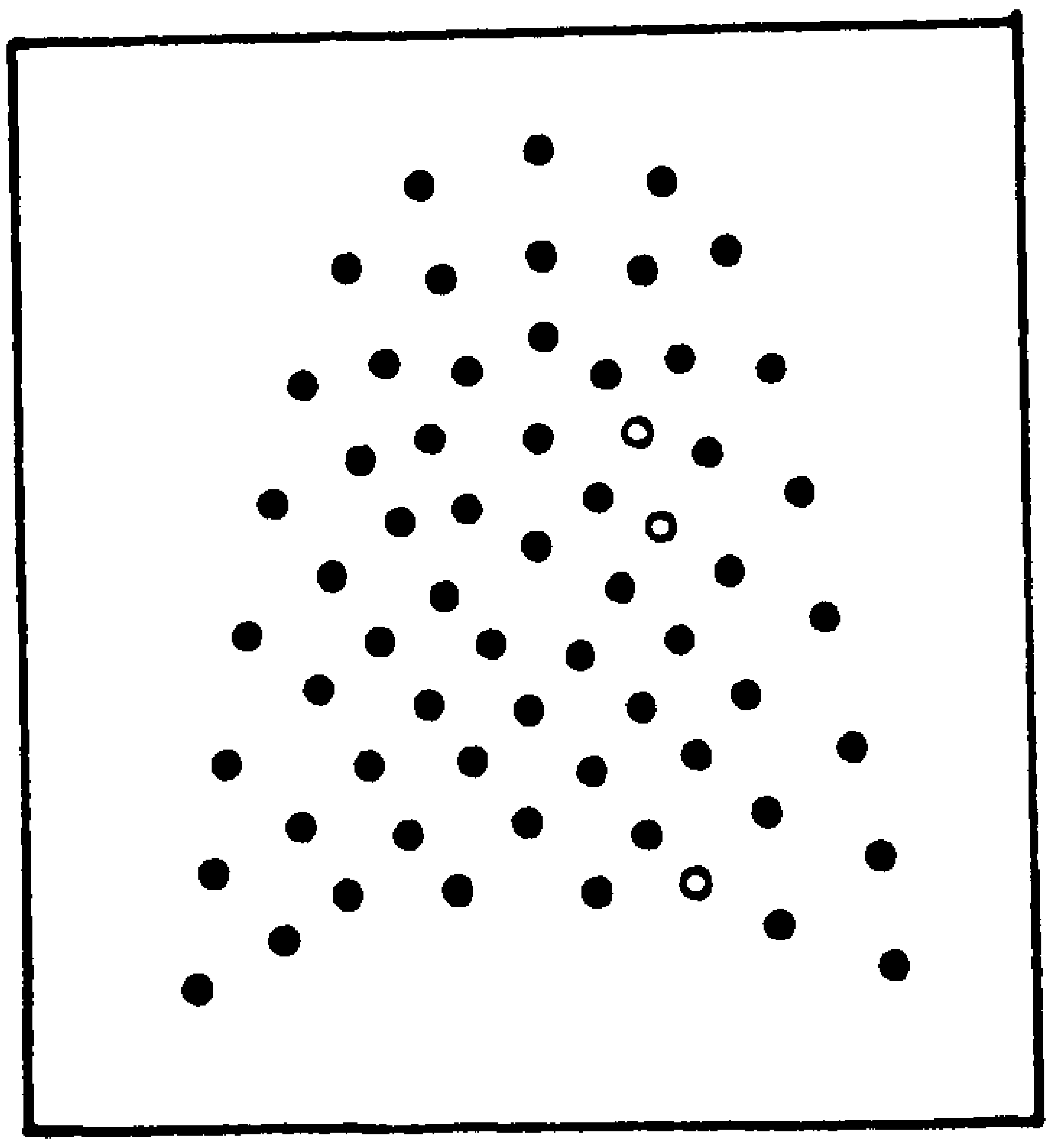
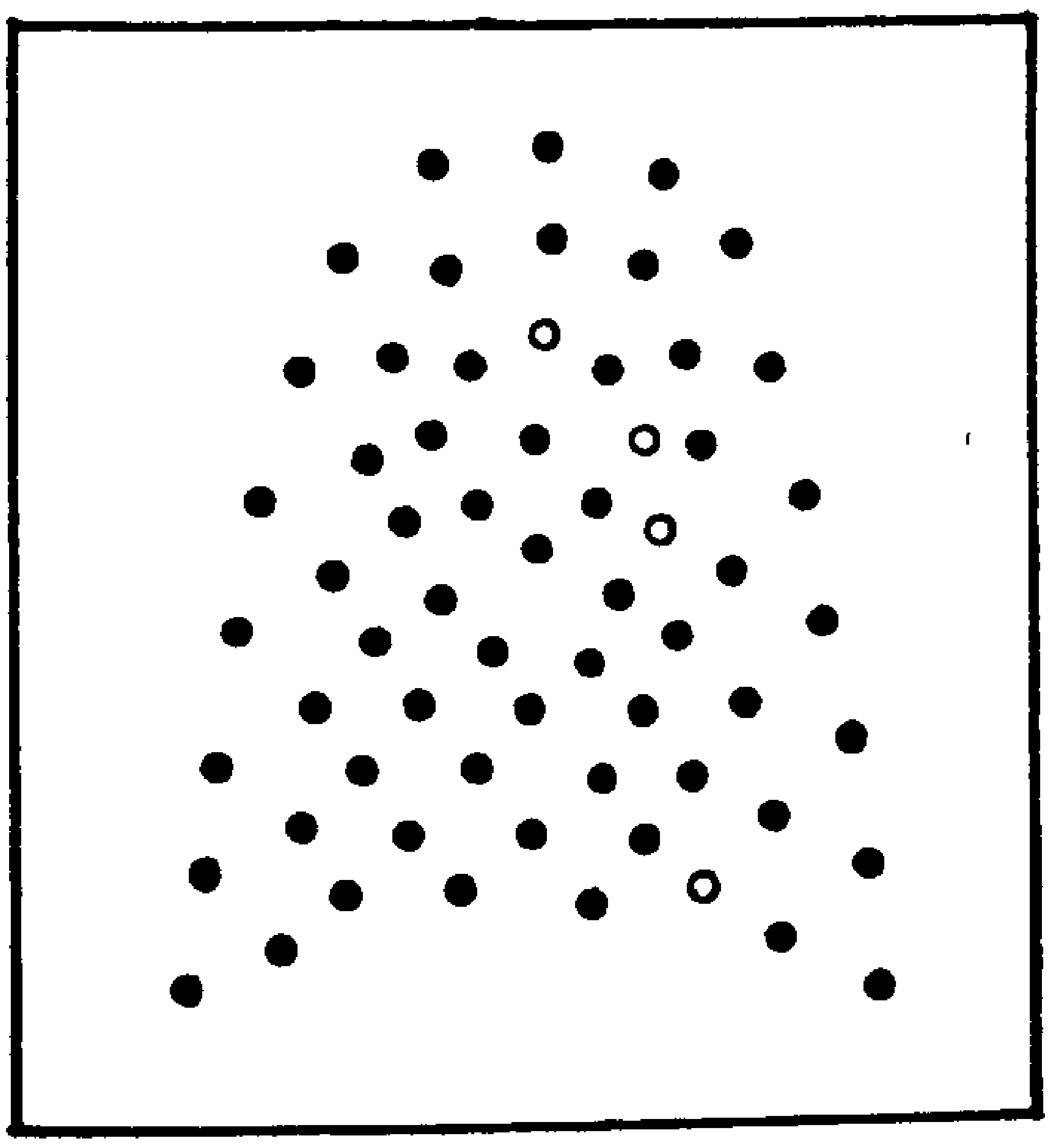
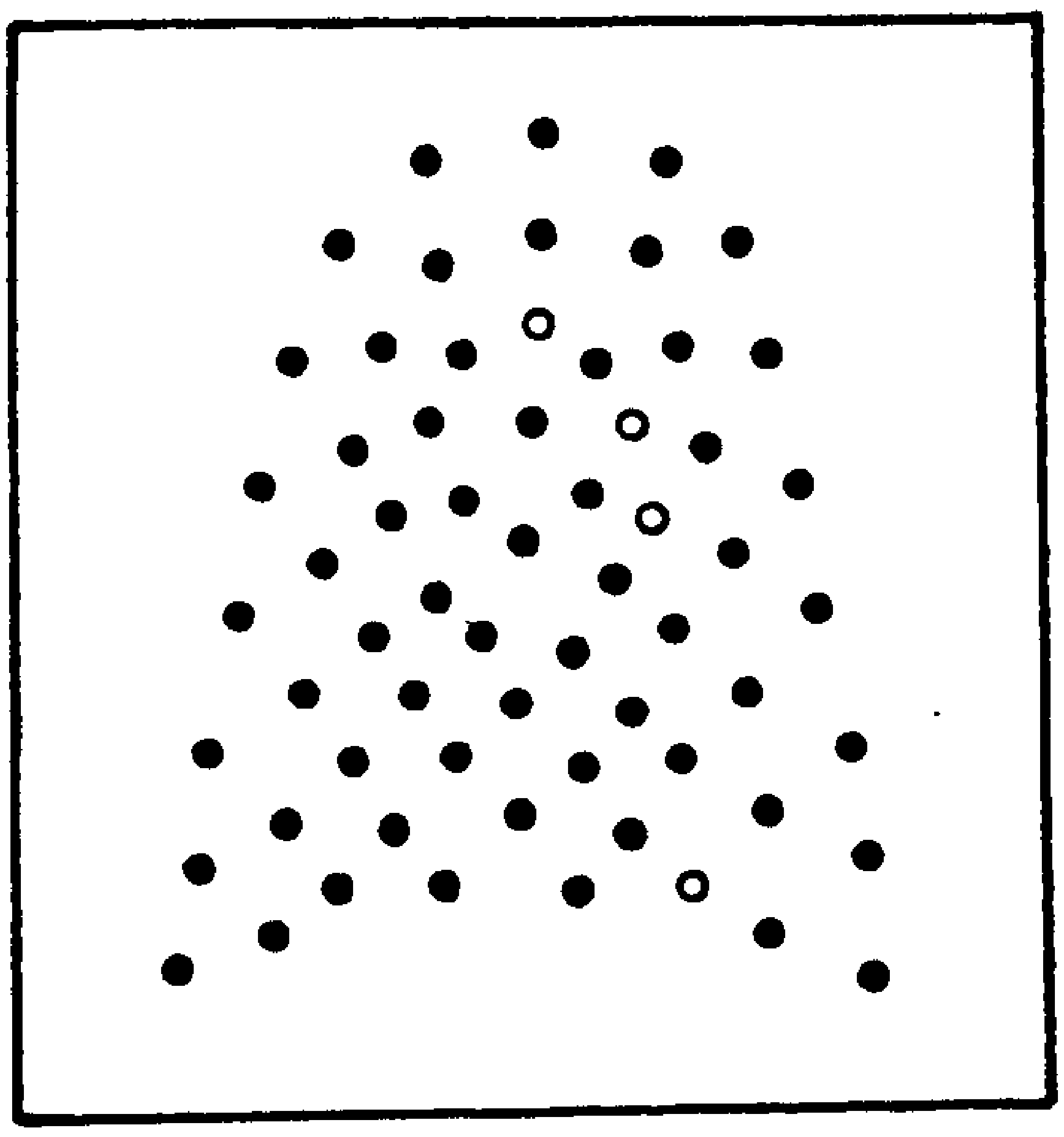


Fig. (5.7) cont'd

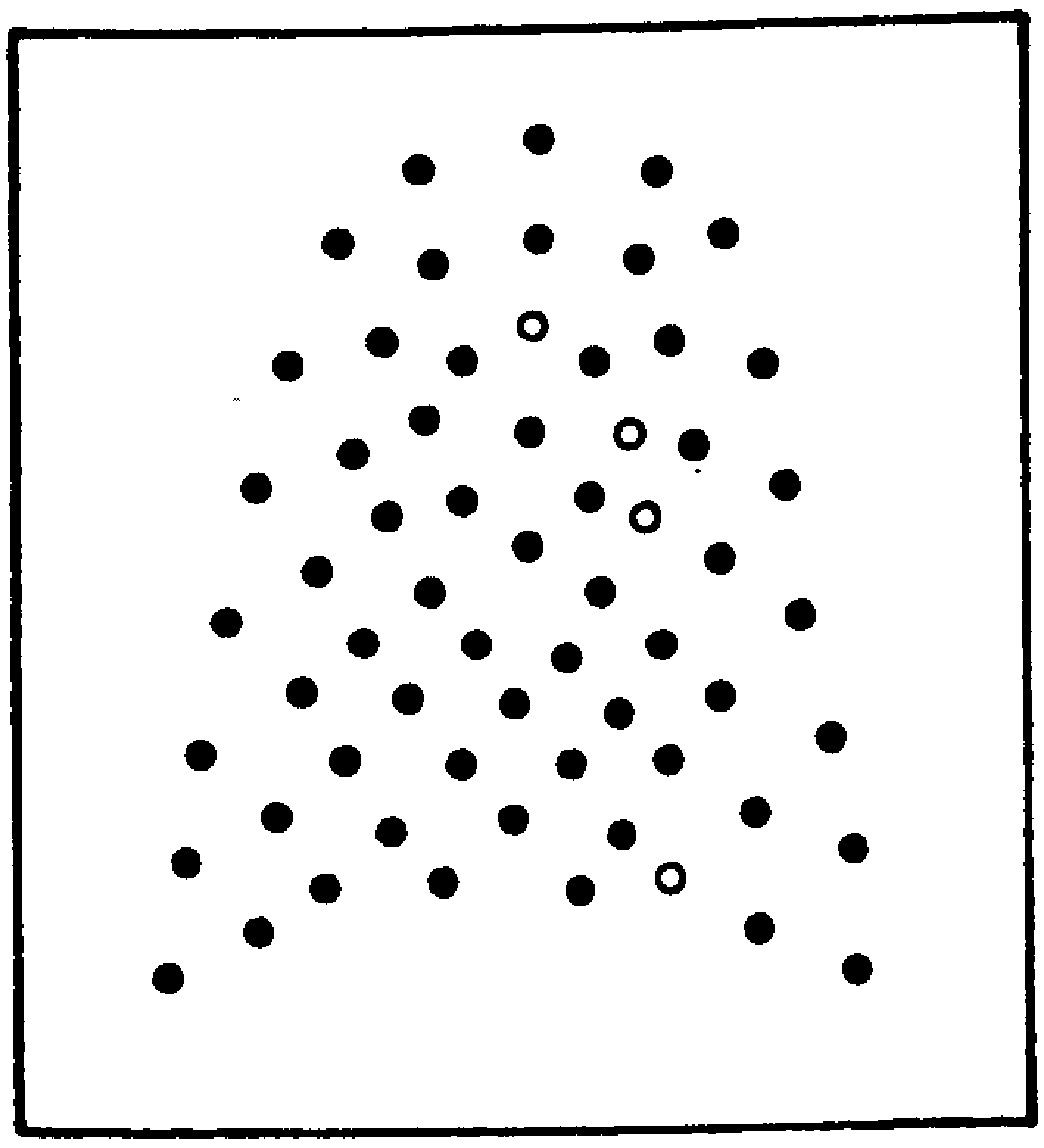
31



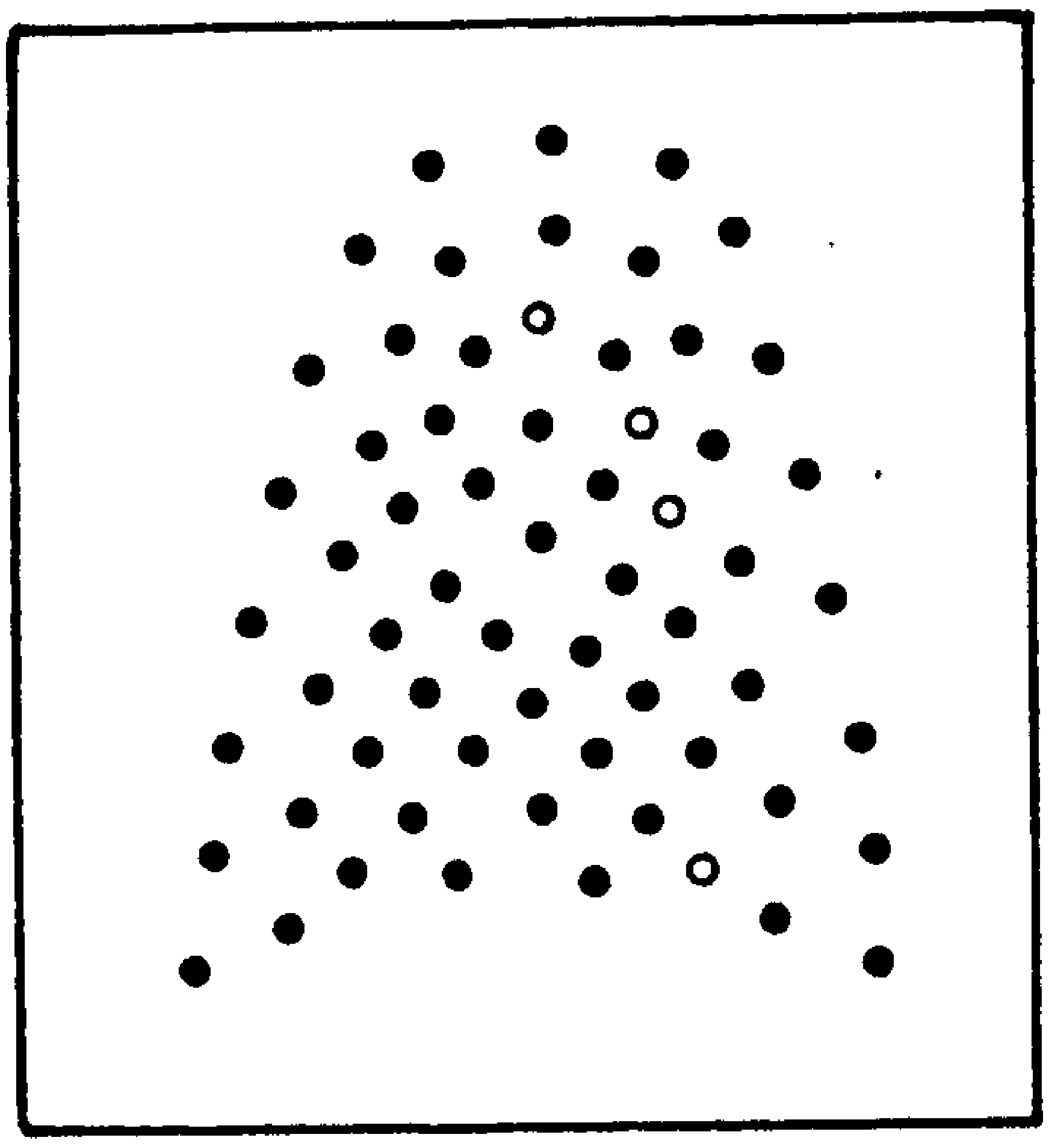
32



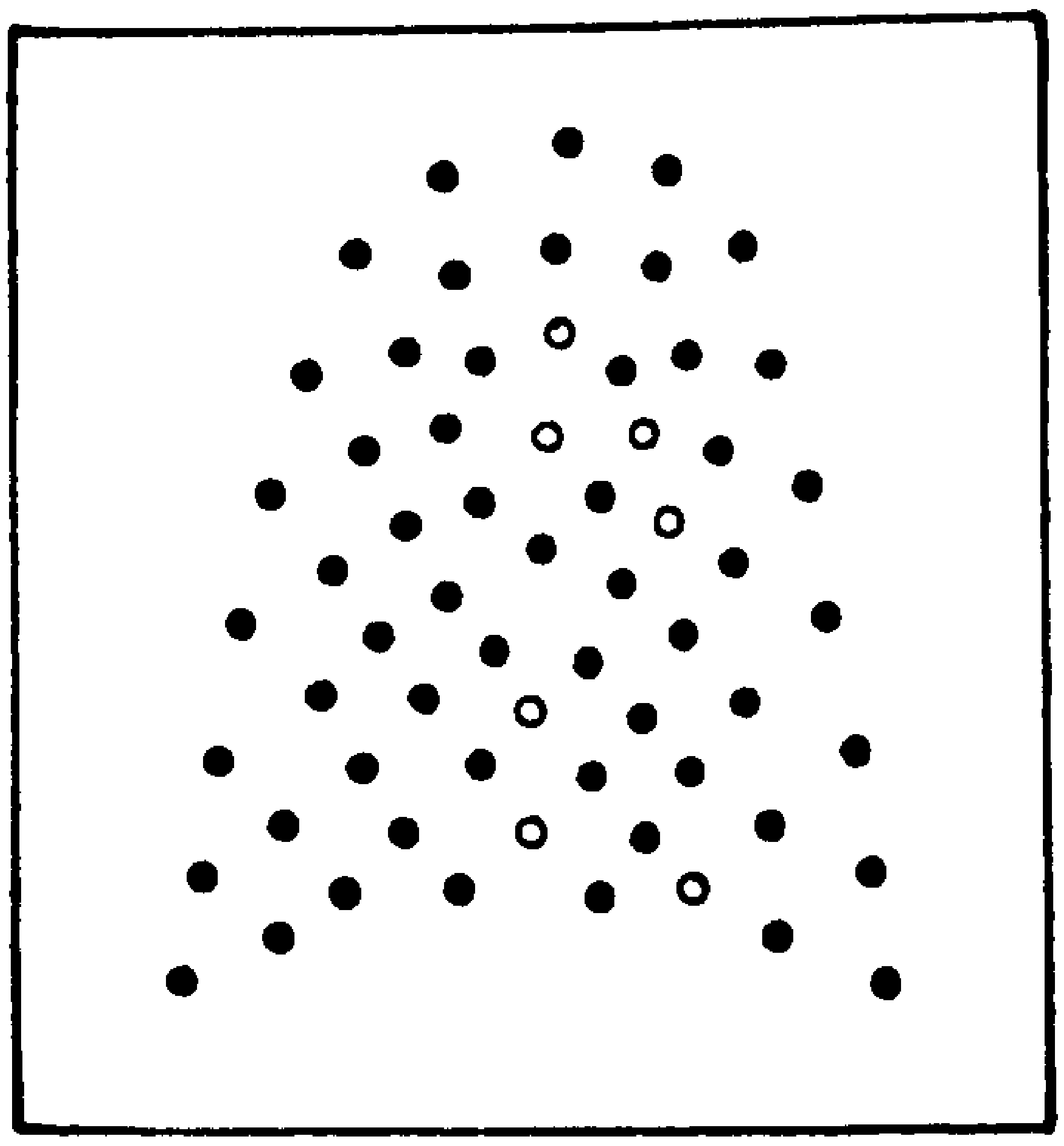
33



34



35



36

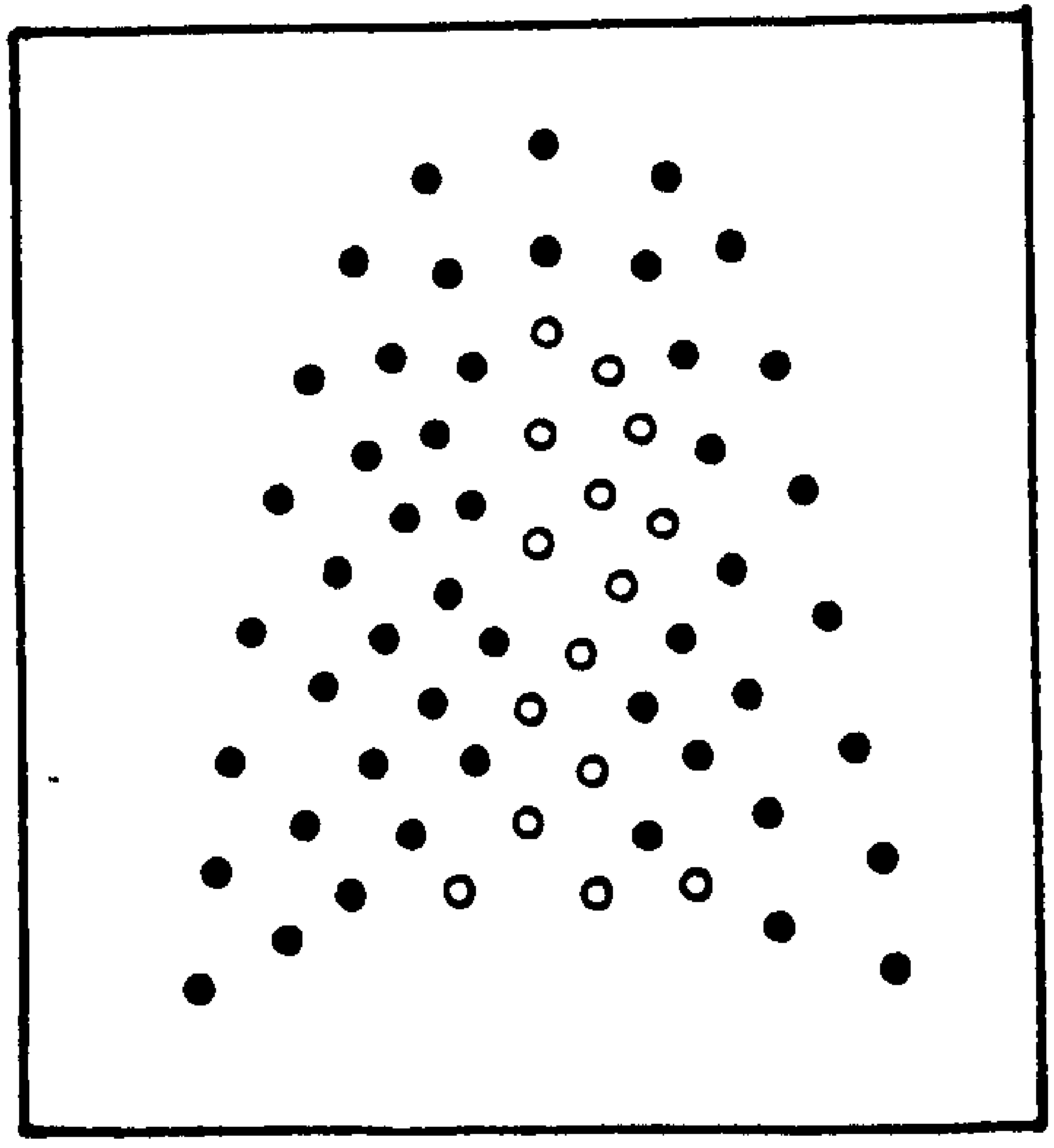
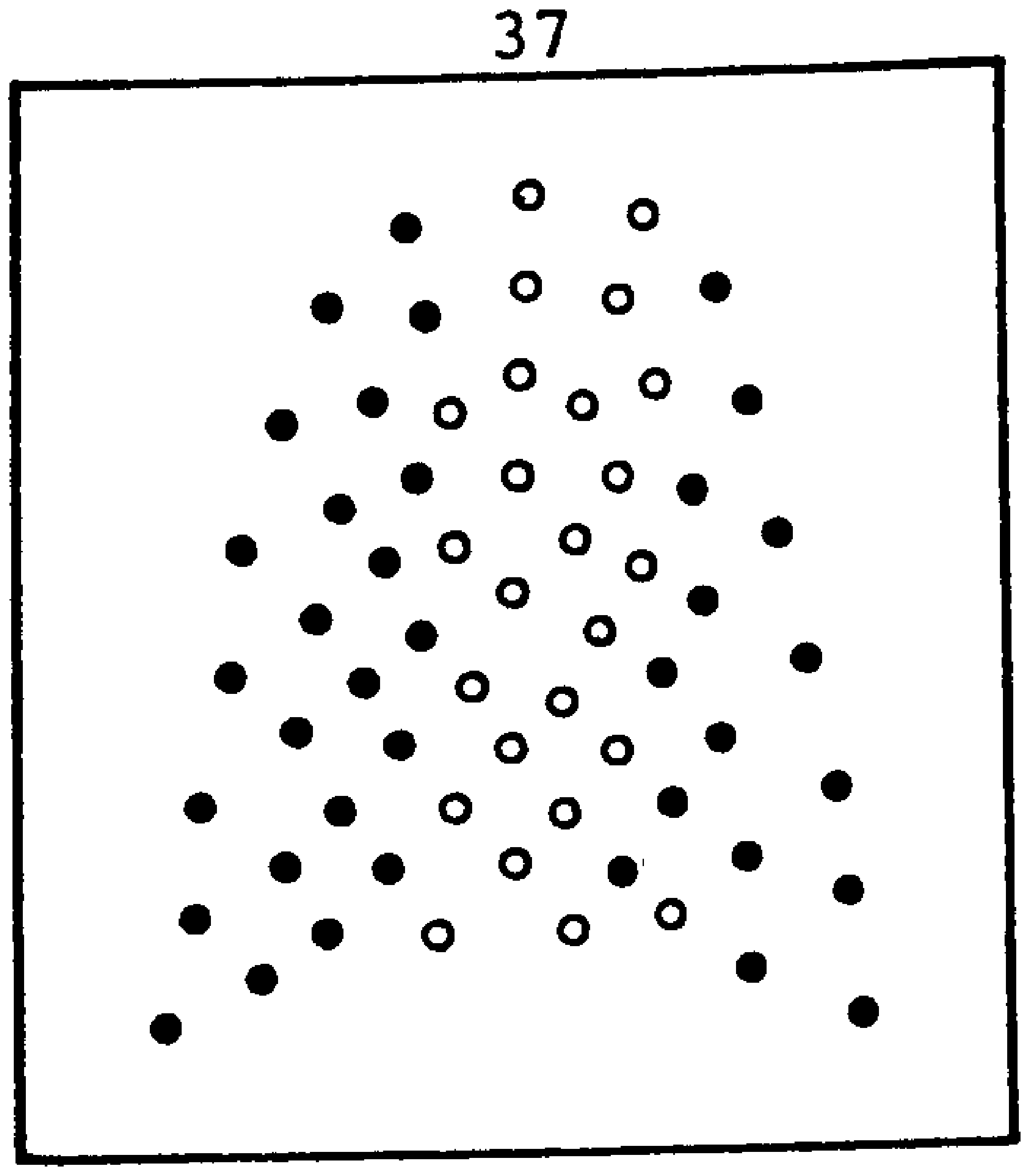


Fig. (5.7) cont'd

37



38

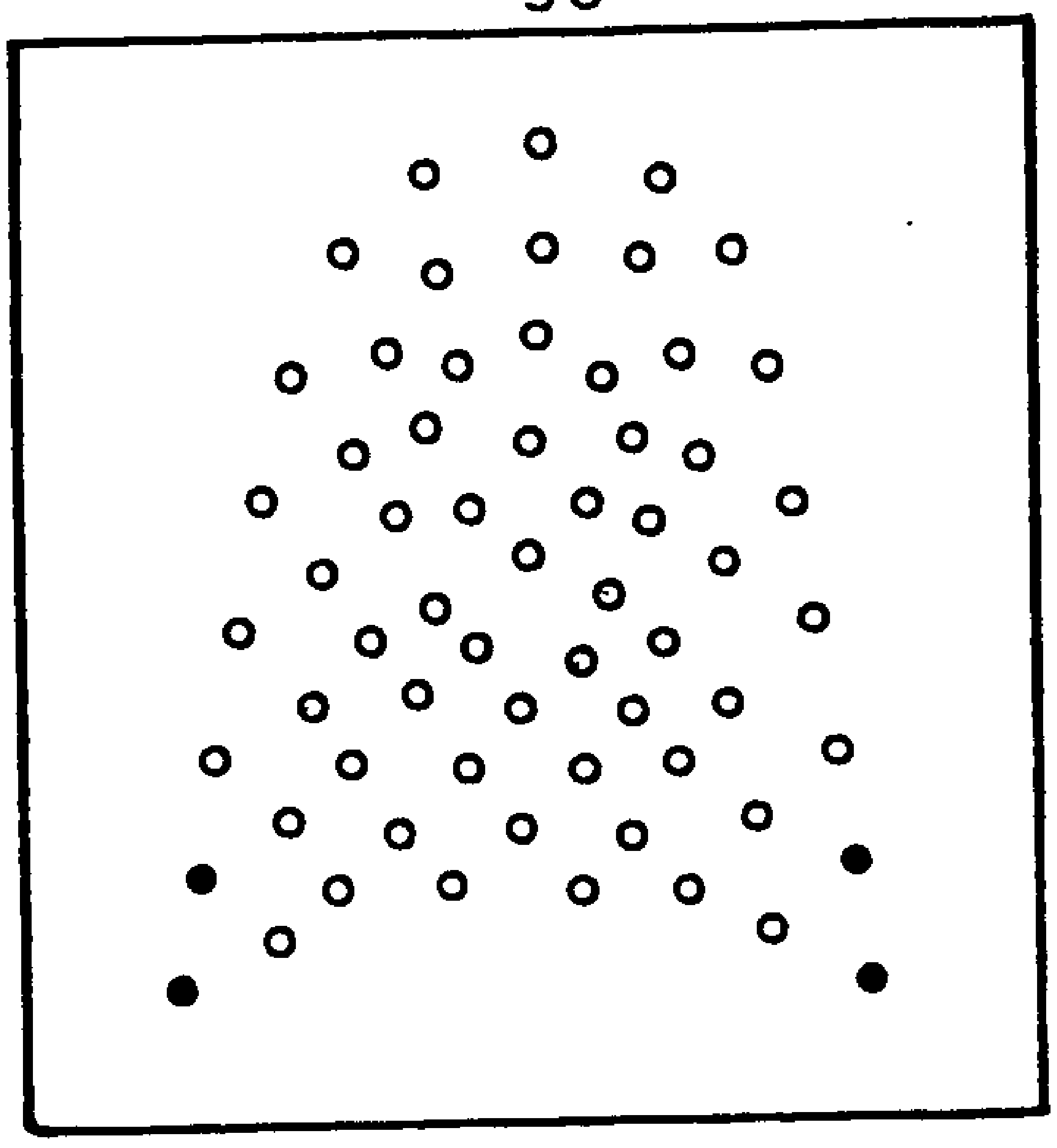


Fig. (5.8) Superimposed graph showing the number of contacts as a function of time (i.e. sample number) for word-medial [s] and [ss] in the environments of /ʕasa/ and /ʕassa/, respectively.

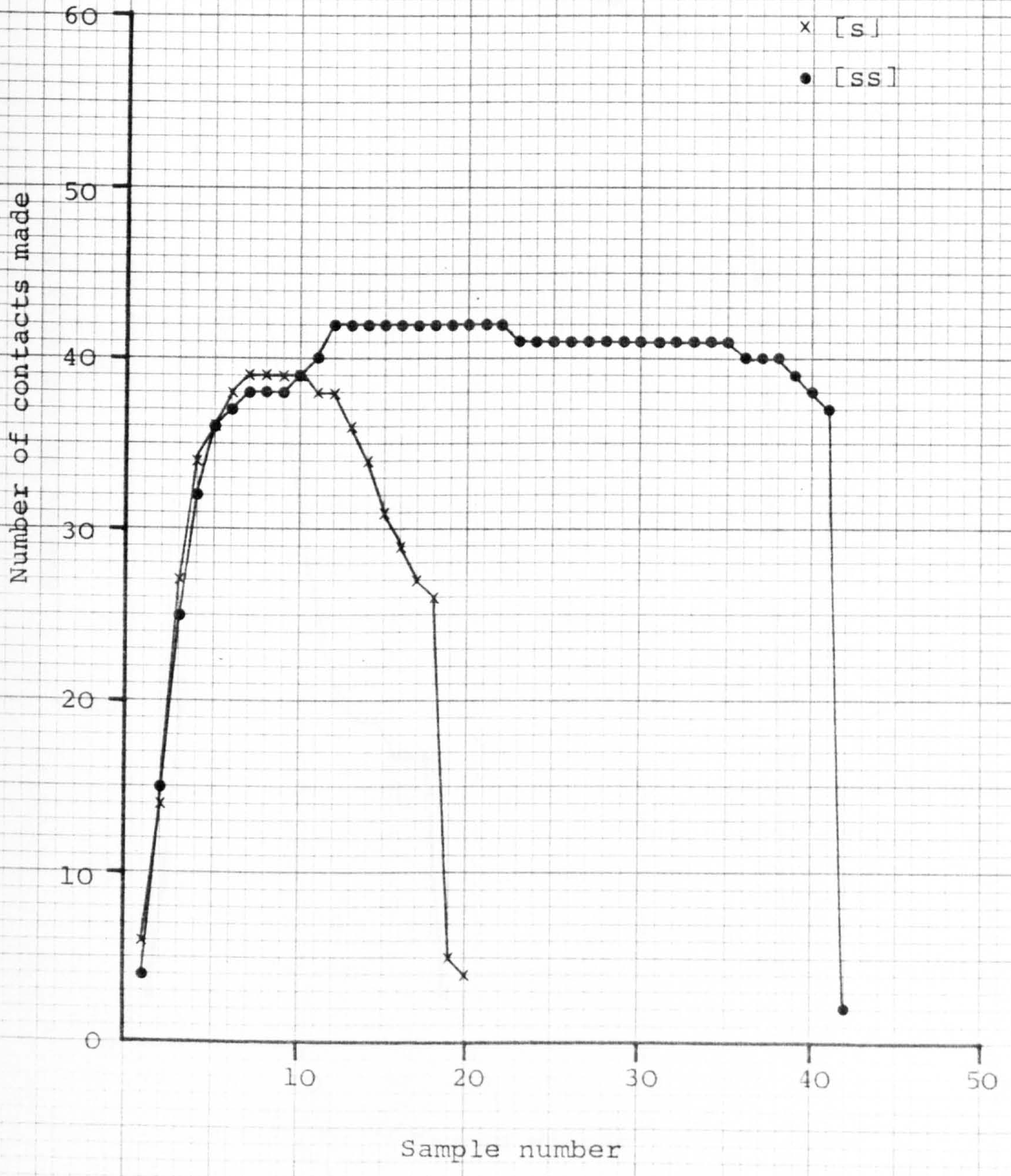
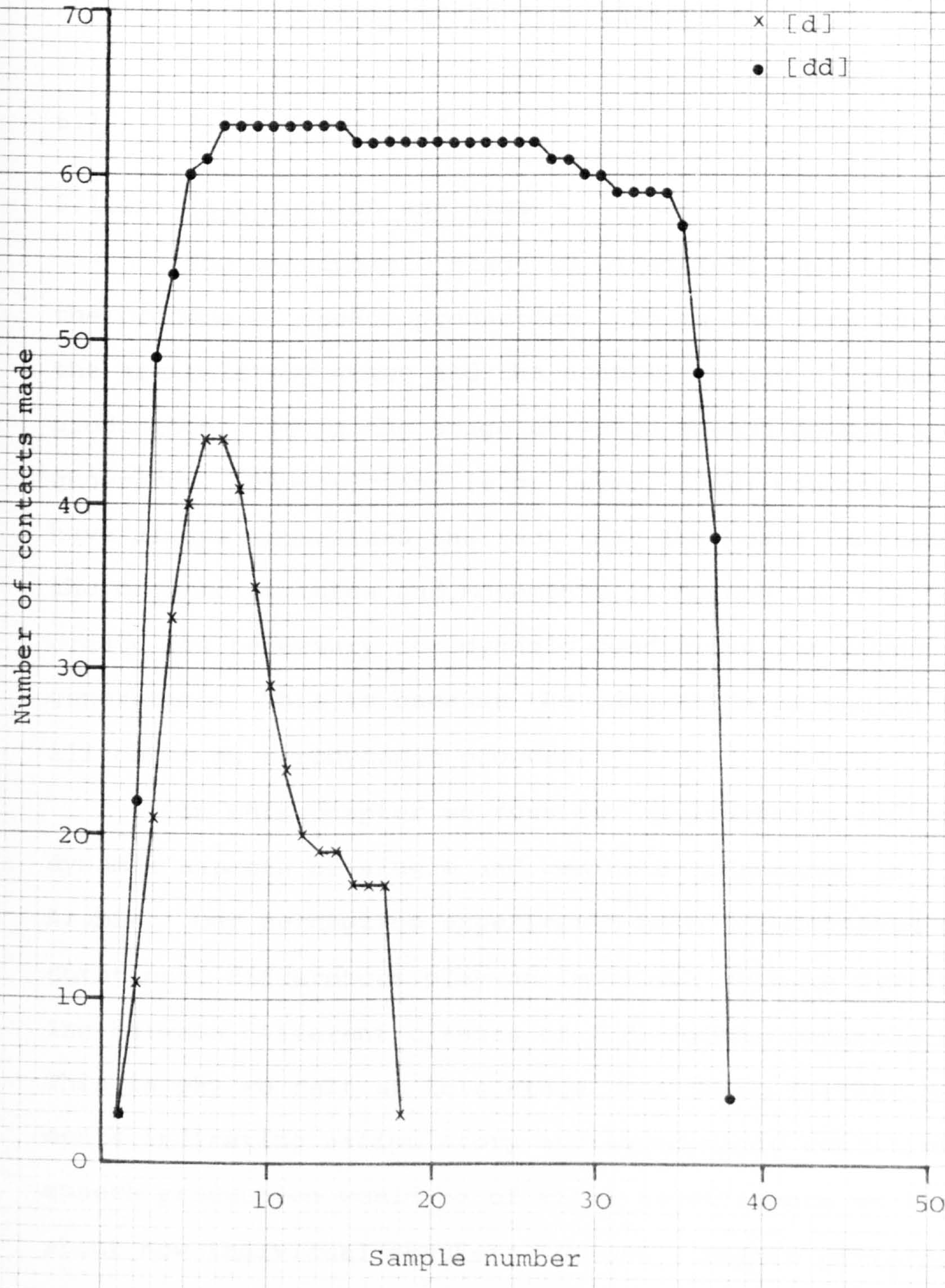


Fig. (5.9) Superimposed graph showing the number of contacts as a function of time (i.e. sample number) for word-medial [d] and [dd] in the environments of /ξada/ and /ξadda/, respectively.



CHAPTER SIX

ARTICULATORY AND AERODYNAMIC STUDY OF GEMINATE CONSONANTS IN I.C. ARABIC

6.1 Pilot Experiment

6.1.1 Introduction and Aims

The aerodynamic stage is of extreme importance in the process of speech production. It is the link between the articulatory stage and the acoustic stage of speech production. It is the stage at which speech sounds are generated. Phoneticians (e.g. Laver, 1968; Catford, 1977) have always emphasized the fact that the basic function of the organic postures and movements in speech is only to create the necessary aerodynamic conditions for sound-generation. This is because the generation of speech is in all cases an aerodynamic process.

In this chapter we consider articulatory and aerodynamic aspects of single and geminate consonants in I.C. Arabic. Our perception experiments have demonstrated that durational differences play an important role in distinguishing between different classes of I.C. Arabic consonants. This is why we felt at this stage that doing further experiments indicating articulatory and aerodynamic conditions of speech production would be of vital significance to learn about how individual speakers achieve acoustic patterns. We can then ask whether there is more invariance of articulation than acoustics and whether there are different options. These experiments may confirm the fact that temporal differences found in our acoustic experiment and in our experiments of synthetic speech correspond to articulatory and aerodynamic differences between single and geminate consonantal segments in I.C. Arabic.

The aim was to compare the articulatory and aerodynamic findings obtained in the present experiment with the acoustic findings reported in Chapter Three of this study. To assist this comparison, it was important to find out whether the acoustic differences in duration stated in that experiment show trends consistent with those found in this experiment. Mingograph traces of the recorded data were produced. This may also help us to find out whether durational differences are affected by the type of the recordings made. That is to say, whether recordings made in the echo-proof recording studio and those made in the laboratory with a mask on face and a tube inserted through the nose down into the pharynx would yield different results.

One of the principal objectives of research in speech is to understand the mechanism underlying the control of the speech-generating system. It is clear that consonant sounds depend greatly on the interaction of articulator movements and airflow patterns. Several types of experimental techniques have been constructed to examine the nature of this process. The acoustic aspects of consonants have been largely investigated through spectrographic procedures. Other techniques such as palatography, cineradiography and electromyography have been exploited to study the articulatory movements necessary for consonant formation, see section 5.2.

In our next section we shall refer to some previous studies that have primarily dealt with the domain of the aerodynamics of speech production. We shall discuss their main interpretations inasmuch as consonants are concerned.

6.2 Articulatory and Aerodynamic Studies:- A Critical Review

It is a well-known principle that air flows from a region of higher pressure to another of lower pressure. Investigators of speech production (e.g. Lubker, 1969; Scully, 1969; Warren, 1976) believe that for normal phonation occurrence during the expiratory phase used during speech, pulmonic pressure must be greater than atmospheric pressure. They suggest that alterations in the configuration and size of the air-passage immensely affect airflow, particularly when the respiratory tract is used for articulation. Warren (1976), for instance, pointed out that "one of the main differences between the airflow patterns in breathing and those in speech results from an orifice type of airflow which develops as constrictions form within the vocal tract during phonation." [p.108]. This type of airflow, he added, may exist at the glottis, the velopharyngeal orifice, and at various points along the oral air-passage.

The voice sound is generated by making airflow pass through the glottis when the vocal folds are suitably adducted. The vibratory movement of the vocal folds is, to some extent, the outcome of aerodynamic and muscular forces which interact in an intricate way. Warren (op.cit.) remarks that the aerodynamic forces which are exerted on the vocal folds include "subglottal air pressure, the Bernoulli force produced by negative pressure created transglottally by high-velocity airflow, and supraglottal pressure produced by articulatory constructions in the upper airway." [p.119]. This may give rise to the fact that in the case of velocity increase, pressure must consequently decrease, and that pressure at the glottal constriction may fall below atmospheric as the vocal folds adduct. Hence, van den Berg et al. (1957) commented that glottal tissue resistance increases as the area of constriction decreases, thus allowing the Bernoulli force to increase only up to a certain point. They stated that "the Bernoulli effect of the air escaping through the narrow glottis is also very important. By accounting for this effect, which causes a negative pressure in the glottis, it has been possible to explain the onset of the vibration of the vocal folds at a breathy attack. The vibrations then start with an inward movement of the vocal folds, caused by the sucking Bernoulli effect." [p.626].[1]

In discussing the effect of the 'myoelastic-aerodynamic' theory of speech production, van den Berg (1958) pointed out that the functioning of the 'glottis-generator' was to be realized solely on the basis of this theory. He stated that the fundamental frequency of this type of generator

---

[1] For further details see van den Berg et al. (1957), p.631 under "Negative Pressure in the Glottis by the Bernoulli Effect".

was equal to the vibrations of the vocal folds, and claimed that this frequency depended on several factors which could help in calculating the frequency. The factors suggested are: "(1) the effective mass of the vibrating part of the vocal folds; (2) the effective tension in the vibrating part of the vocal folds; (3) the effective area of the glottis during the cycle, which determines the effective resistance of the glottis and the effective value of the Bernoulli effect in the glottis; (4) the effective value of the subglottic pressure; (5) the damping of the vocal folds." [p.239]. He concluded that the closure of the glottis is brought about by three basic factors; the decrease of the subglottic pressure due to the escape of subglottic air, the tension of the vocal folds which causes the glottal closure, and the sucking effect of the escaping air.

Earlier, van den Berg (1956) proposed that a knowledge of the subglottic pressure would be of great importance because the energy required for phonation and produced via the respiratory muscles is proportional to the product of the mean subglottic pressure and the mean airflow, "at least in a first approximation." [p.1]. Elsewhere, van den Berg et al. (1957) added that a knowledge of the air resistance of the larynx as a function of the subglottic pressure, the dimensions of the glottis and the pressure in the laryngeal ventricle would also be of value so as to estimate the human glottis generator properties. They succeeded in measuring this resistance "under conditions of constant flow by means of a model of the larynx, prepared from a cast made from a normal fresh human larynx." [p.626]. They also measured the sucking Bernoulli effect with the model of the larynx they designed. They defined the resistance (R) of the model using the following formula:

$$R = \frac{\text{subglottic pressure (Ps)}}{\text{volume velocity (V)}} \text{ dyne sec/cm}^5.$$

where (R) is composed of three elements, thus:

$R = R_1 + R_2 + R_3$

$R_1 = (Ps - P_1)/V$ is the resistance at the entry of the glottis,

$R_2 = (P_1 - P_4)/V$ is the resistance in the glottis, and

$R_3 = P_4/V$ is the resistance at the outlet of the glottis.

Van den Berg et al. (op.cit.) propounded the view that turbulence is not restricted to the glottis; this is due to the fact that 'eddies' are generated before the air is forced into the glottis. Their measurements indicated that "the critical volume velocity for this formation was smaller at the entry of the glottis than in the glottis." [p.627]. This results in a larger resistance at the entry of the glottis (R_1); while enlargement of the resistance in the glottis (R_2) is accomplished when 'eddies' are generated at the entry of the glottis where they are carried by the flow into the glottis. They also pointed out that there is an increase in the resistance at the outlet of the glottis (R_3) when 'eddies' are generated at the outlet of the glottis. This, they contended, is due to the partial waste of the kinetic energy of the air "by the formation of eddies in the laryngeal ventricle." [p.628].

Recently, Titze (1980) has quantified and tested with mathematical models the myoelastic-aerodynamic theory of voicing. He claimed that the models suggested that vocal fold oscillation is produced as a result of asymmetric forcing functions over closing and opening portions of the glottal cycle. The ranges of oscillation increased among various models as more freedom in the simulated tissue movement was incorporated. For him the vertical motions of the vocal folds tissues were of particular significance in initiating and maintaining oscillation. This is because these motions "facilitate coupling of aerodynamic energy into the tissues and allow tissue deformations under conditions of incompressibility." [p.495]. He regarded the control of fundamental frequency of oscillation (FO) as basically myoelastic, partly because of deliberate or reflex adjustments of laryngeal muscles, and partly because of nonlinear tissue strain over the vibrational cycle. He added that "This places limits on the control of FO by subglottal pressure, and forces such control to be inseparably connected with vibrational amplitude, or less directly, with vocal intensity." [loc.cit.].

Rothenberg (1968) stated that resistance at the vocal folds varies from less than 1.0 cm H_2O/L/sec during quiet breathing to as much as 100 cm H_2O/L/sec during voicing. For example, the difference between subglottal and supraglottal pressures during vowel articulation is approximately between 4.0 to 8.0 cm H_2O, and between 1.0 to 4.0 cm H_2O during voiced consonant articulation. He added that during phonation, the oral air-passage produces resistance of less than 1.0 cm H_2O/L/sec for certain vowels to an infinite resistance or complete obstruction for plosives. According to Warren (1976) the larger differences between subglottal and supraglottal pressures during vowel production are due to a lower resistance in the upper air-passage. This is because, for vowels, the pressure above the glottis is only slightly above atmospheric pressure, so that most of the subglottal pressure creates the pressure drop across the glottis. He also stated that voicing entails the passing of airflow across the vocal folds where a differential pressure between the subglottal and supraglottal cavities must occur. These conditions do not always obtain in the case of the production of plosive consonants. Kent and Moll (1969), on the other hand, demonstrated that voiced plosives are produced with larger supraglottal volumes than their voiceless counterparts. They argued that the increase in supraglottal volume is at least a partial explanation for intraoral pressure differences among cognates.

It is worth mentioning here that Chomsky and Halle (1968) restricted the 'tense-lax' feature entirely to supraglottal articulations, and did not consider it to be associated with tensing of the vocal folds in any way. They claimed that in the case of voiced plosives the walls of the pharynx are necessarily lax, and that it is this laxing which causes the enlargement of the vocal tract which is required during the production of voiced plosives. The enlargement of the pharynx during voiced plosives is a well-known phenomenon.

One investigator states that "this enlargement is not due to the increase in oral pressure resulting in the lax walls of the pharynx being pushed back; nor is it generally held that in voiceless stops there is no such enlargement because the tense walls of the pharynx actively resist it." (Ladefoged, 1971, p.96).

The conclusions reached by Chomsky and Halle (op.cit.) were based on data reported by Perkell (1965, 1967). The indirect observations stated by Lisker and Abramson (1967) and the cineradiographic studies performed by Kent and Moll (op.cit.) have shown opposite conclusions to those of Chomsky and Halle. They stated that the enlargement of the vocal tract during voiced stops is an active muscular process, and there is no increase in the tension of the pharyngeal muscles during voiceless stops. (Ladefoged, ibid., p.97).

In a study by Slis (1970) it has been shown that if the intraoral pressure rises, the pressure drop over the glottis, i.e. subglottal pressure minus intraoral pressure, will decrease. Consequently, this results in a decrease of the air-stream through the glottis. Slis remarked that this decrease of the air-stream would take place rather quickly with voiceless consonants in which the stream of air would soon be too small to maintain the vibrations of the vocal folds. With voiced consonants the pressure drop would remain 'sufficiently high' to keep the vocal folds in vibration. He showed that the air pressure in the mouth-throat cavity rises quickly owing to the volume decrease and low glottal resistance, i.e. when the glottis is open.

Following his previous studies and that of van den Berg (1956), Slis (op.cit.) obtained his direct measurements of the intraoral pressure by means of an open rubber tube which was pushed down the nose into the pharynx. The tube was connected with a pressure transducer outside his subject. He measured the subglottal pressure indirectly via a small rubber balloon which was introduced into the subject's oesophagus. Through the nose, the balloon was connected with a pressure transducer outside the subject by means of a rubber tube. His results illustrated that "peak values of the subglottal pressure seem to be independent of the consonants except for /h/ during which a pressure drop could be observed." [p.200]. The intraoral pressure, both in normal speech and in shouting, was shown to be considerably higher during voiceless consonants than during voiced ones; while in whispered speech these differences were rare. Slis noticed that in normal speech the intraoral pressure indicated a sudden rise followed by a more gradual rise. He believed that the two rises were presumably due to the closing of the mouth and to the 'inflation' of the pharynx, respectively.

Depending on his final results, Slis (op.cit.) reached the conclusion that "differences in intraoral pressure do not have their origin in differences in subglottal pressure; on the contrary, during shouting we even observed a tendency to a higher subglottal pressure in voiced consonants than in voiceless consonants, while the intraoral pressure was lower." [p.200]. The explanation proposed for this was that there was a higher resistance in the vibrating glottis. The glottis proved to be more open during voiceless consonants than during voiced ones. See table 6.1 below.

	Normal speech	Shouting	Whispering	Number of measurements n
Voiceless plosives	60	130	85	6
Voiceless fricatives	60	125	70	6
Voiced plosives	35	65	70	4
Voiced fricatives	50	95	60	4
Glides, nasals, liquids	20	20	40	12
/h/	5	5	5	2

Table (6.1) Peak values of the intraoral pressure (in mmH_2O) during different classes of consonants. The number given is the mean of n measurements.

(After Slis, 1970)

Much of the available published data on intraoral pressure have, generally speaking, been derived from recordings of isolated monosyllabic and/or bisyllabic nonsense words (e.g. Malécot, 1966). Subtelny et al. (1966) have shown that the factor of overall vocal effort, viz. loudness, and syllable rate, greatly affect intraoral pressure. The effect of vocal effort has been shown to be as large as that attributable to any linguistic contrast. (Lisker, 1970). In discussing the effects of stress in nonsense words Scully (1971), on the other hand, has pointed out that from these words "it was not possible to draw firm conclusions about the effect of stress upon the aerodynamic variables." [p.193]. An examination of the obtainable traces has indicated that the tongue constriction flow resistance (Rc) reached a lower minimum and that the cross-section of the tongue constriction (Ac) reached a greater maximum during the stressed vowel of a word. The results showed that a similar value of average airflow had been observed during both stressed and unstressed syllables. A greater degree of opening of the tongue constriction had been noticed in the stressed syllables. Previously, Arkebauer (1964) reported data relative to changes in air pressure with changes in consonant intensity level. His findings were that as the vocal intensity of a consonant-vowel syllable increased, the intraoral air pressure associated with the consonant increased concomitantly. Other investigators (Kunze, 1964; Ladefoged and McKinney, 1963; Isshiki and Ringel, 1964) have demonstrated that increases in the intensity level of the voice are accompanied by increases in subglottic air pressure, and that air pressures generated within the vocal tract increase as the level of speech is raised.

In two subsequent investigations carried out on ten children and ten adults, Arkebauer et al. (1967) have observed that the mean peak intraoral air pressures were substantially higher for voiceless consonants than for their voiced counterparts, irrespective of their phonetic context, and that "the absolute differences between plosive cognates were greater than those between fricative cognates." [p.199]. For voiceless consonants, they showed that plosives produced higher mean peak intraoral air pressures than fricatives in all phonetic contexts. However, within both the voiceless plosive and fricative classifications "the consonant rank seems to be dependent upon phonetic context, except for /f/ which consistently exhibits the lowest pressure for the voiceless consonants." [pp.199-200]. The mean intraoral air pressure was found to diminish as the consonants were produced in three different contexts, namely in inter-, pre-, and post-vocalic contexts, respectively.

For their adult group, the results indicated that the peak intraoral pressures were consistently higher for voiceless consonants than for their voiced partners, regardless of utterance rate or intensity level. It was found that with voiceless consonants mean peak intraoral air pressures were greater for plosives than for fricatives under all experimental conditions, and that greater intraoral pressures were always associated with consonants occurring in intervocalic positions than in pre- or post-vocalic positions. Higher pressure magnitudes were found to accompany consonants in prevocalic context than in postvocalic context. Arkebauer and his colleagues (op.cit.) claimed that their findings were in agreement with the data of Hixon (1966) who had presented evidence that peak intraoral air pressures increase as a function of speech intensity, but were not in agreement with those of Black (1950) and Malécot (1955). Black had found that peak intraoral air pressures diminish as consonants are produced in pre-, inter-, and post-vocalic positions; whereas Malécot had found that intervocalic consonants are accompanied by greater peak intraoral air pressures than their pre- or post-vocalic consonants.

Hixon's research (1966) was designed to investigate "simultaneous variations in intraoral pressure, oral airflow rate, and the area of the maximum oral constriction associated with specified conditions of turbulent noise production for speech." [p.168]. He measured intraoral air pressure, rate of oral airflow and consonant sound pressure level at the peak consonant noise level within each consonant articulation chosen for his study. The simultaneous values of air pressure and airflow were used to estimate the maximum oral constriction area at each measured point in time within each consonant articulation. The measurements were computed by using a formula similar to that reported by Warren and DuBois (1964) for measuring the area of the velopharyngeal orifice. Nevertheless, Hixon (ibid.) assumed that it was possible to approximate the area of the maximum oral constriction by neglecting the empirically determined correction factor which was inserted in Warren and DuBois's equation[1] to account for differences between the actual area and the theoretical area which result from turbulent, nonuniform and rotational flow. He neglected this correction factor because he thought it was unnecessary "since the interest in the present study was in relative rather than absolute areas." [p.174].

---

[1] For further details pertaining to this equation see section 6.3.3.3.

Hixon's findings suggested that changes in sound pressure level for voiceless fricatives were accompanied by similar trends of change in intraoral air pressure. The results indicated that there was a rapid increase in airflow rate during the initial part of the consonant. This rapid increase remained relatively constant for a while and was then followed by a noticeable rise in airflow rate. The rate of change in constriction area was observed to be considerably greater during the initial and terminal parts of the production of the consonant than in that part preceding the peak sound pressure level. The results also indicated that air pressure, airflow and constriction area were not affected by utterance rate.

Black (1950) selected eight different consonants for his study to investigate the relative amounts of air pressure in the mouth that accompany the production of each consonant. The selected consonants represented voiced-voiceless and plosive-fricative contrasts. They were arranged in three positions, viz. initial, medial and final, with three different vowels. In a medial position, the consonant was preceded and followed by the same vowel sound. His results showed that the consonant under investigation, in almost all the examined comparisons, was accompanied by a diminishing pressure as the consonant receded in the end. Final consonants were found to be spoken with less pressure than initial ones.

Malécot (1955) reported that the average of the peak pressures for a given voiceless consonant is regularly greater than for a voiced consonant, and that voiceless consonants are attributed greater 'force of articulation' than their voiced cognates in all but one pair, namely [s] : [z]. Later, Malécot (1966) made investigations "to study individually the various air-pressure-pulse parameters that may be involved in feedback." [p.66]. He was primarily concerned with pretonic intervocalic single plosives in American English used in nonsense syllables, such as /apa/ and /aba/. His results suggested that all three parameters involved in the study, i.e. peak pressure, pressure impulse and duration, were shown to be distinctive, and a statistical superiority was given to the first two parameters (peak pressure and pressure impulse) as compared to the third parameter (duration).

Further, Malécot (1970) demonstrated that the 'fortis-lenis' feature of consonants "is primarily a synesthetic interpretation of magnitudes of intrabuccal air pressure and is conveyed variously in different contexts by the durations of the consonant closure and of the preceding vowel." [p.1588]. This fact was previously observed by Malécot (1966). He suggested that 'force of articulation' is a significant attribute of consonants and it is a determinant factor in the 'fortis-lenis' opposition particularly in pairs such as [p] : [b] and [s] : [z]. He also considered it as a linguistic reality.

Malécot's collective results on 'force of articulation' have been critically reviewed by Ashby (1977) who commented that there is 'notoriously' little agreement among phoneticians over the meaning of 'force of articulation'. Fortis sounds are often characterized by greater 'force', 'effort' or 'tension' of articulation, and greater 'breath force' or 'degree of breath'. More specifically, they have been equated with the following four factors : (Ashby, op.cit., p.74).

1) greater force at the point of articulation;
2) increased muscular tension elsewhere in the supraglottal tract;
3) increased subglottal air pressure;
4) increased intraoral air pressure.

Recent studies have shown that a fortis consonant exhibits a higher intraoral air pressure than its lenis cognate when all other three factors are kept constant.

According to Ashby, Malécot was concerned with the lenis-fortis feature only as it corresponds with the voiced-voiceless contrast, and "makes no mention of the various reported cases of fortis articulation as a separate feature, affecting both voiced and voiceless segments." [p.75]. He pointed out that Malécot's work was not primarily an account of those factors exemplifying the lenis-fortis contrast in either English or French, but an explication of the lenis-fortis concept itself, and that Malécot's conclusions for French were in fact identical with those previously reported for English. Moreover, Ashby criticized Malécot's measurements of pressure curves obtained in his experiments and asked whether 'impulsion' was a valuable measure, or whether it represented "the well-known duration differences in the guise of pressure differences." [p.76].

In conjunction with those experiments designed to quantify the effects on intraoral pressure, intensity, and rate of speaking, Ashby (op.cit.) claimed that Malécot was "very far from having shown that a categorical distinction exists between fortis and lenis types in terms of intraoral pressure." [p.77]. Several extraneous factors such as loudness of speech, prosodic conditions and speaker characteristics might cause differences of pressure at least as great as those stated to separate fortis from lenis. Eventually, Ashby came to the conclusion that Malécot failed to provide sufficient proof that "intraoral pressure, or any other single physical parameter, can provide a contrast categorical enough to serve as a basis for the phonological distinction in question." [p.79]. He also added that it was obvious that 'lenis' and 'fortis' are as much cover terms as are 'voiced' and 'voiceless'.

Haag (1977) looked at the general conclusion reported by Subtelny et al. (op.cit.) that "intraoral pressure and peak rate of oral flow are generally, perhaps linearly, related." [p.514]. His investigation of German plosives was an attempt, based on the speech evidence of only one speaker, to prove the validity of this conclusion. He selected 20 bisyllabic isolated words of the general phonemic structure CVCV which he used in his experiments. Haag found that his results did not confirm the claim put forward by Subtelny et al. (op.cit.), and which is quoted above. His findings showed that the peak airflow for German voiceless plosives was influenced by the degree of opening of the glottis and by the expiratory activity, rather than by the peak oral pressure. He remarked that his findings were in agreement with those of Lisker (1970) and that his data indicated that "peak oral air pressure is positively related to the speaking rate but that the peak oral airflow is inversely related to the speaking rate." [p.37].

On the other hand, the experiments conducted by Kunze (1964) to investigate the validity of obtaining subglottal air pressure by using three different techniques showed that measures of intra-oesophageal pressure did not provide an adequate estimate of subglottal pressures irrespective of the size or location of the balloon used; whereas intra-tracheal pressure provided reliable estimate of subglottal pressure during both sustained phonation and connected speech. The results indicated that the intra-oesophageal pressure drop that accompanied the interruption of phonation with abducted vocal folds and constant pulmonic volume provided a valid estimate of subglottal pressure under sustained phonation conditions. The relationship observed in Kunze's study was found to be consistent with those interpretations of previous findings in respiratory physiology, such as those reported by van den Berg (1956).

Isshiki and Ringel (1964) recorded air volume, flow rate, and voice simultaneously while each of eight subjects read a list of 40 items composed of 20 CV syllables and 20 VC syllables. Their data revealed that the mean flow rate values for the voiceless consonants were greater than those for their voiced cognates. This observation was also reported by Ladefoged (1963), and Isshiki and Ringel (op.cit.) related this to the second of the two factors: "expiratory effort and degree of resistance within the vocal tract." [p.241]. Different flow rate patterns were seen to exist for the plosives, fricatives and vowel-like sounds, respectively. The flow rates for the plosive sounds were substantially greater than those found for either the fricatives or vowel-like sounds. Their results indicated that the flow rate for the fricative sounds increased at a steady but gradual pace until a point was reached beyond which the flow rate decreased. They assumed that "this point may correspond to relief from the anatomical stricture necessary for the production of a fricative sound." [pp.238-239].

According to Isshiki and Ringel (op.cit.) the production of a 'typical' plosive sound is composed of three stages which are chronologically arranged. The three stages are: "(1) an incomplete closure of the oral and naso-pharyngeal apertures and an associated reduction in the rate of airflow; (2) a complete closure of the flow ports and total stoppage of airflow; (3) a rapid opening of the flow ports and a subsequent sudden burst of airflow." [p.236]. They noticed that during the production of plosives rapid and distinct variations took place in the shape of the vocal tract. Therefore, they suggested four factors to affect the flow rate during the explosion stage of a plosive sound production. The suggested four factors are: "(1) the opening characteristics of the oral and velopharyngeal stricture, (2) initial pressure within the vocal cavity behind the point of vocal tract closure, (3) the volume of the cavity in which the pressure is built up, and (4) the pulmonary air supply during the period of explosion." [p.243].

In another study, Isshiki (1965) examined the relationship between the vocal intensity and the flow rate at three different pitch levels. His results demonstrated that the variation of flow rate, as a function of vocal intensity, was greater at a high pitch than that at a low pitch. Considerable variations were noticed within individual responses. The results also showed that on the low pitch phonation, the flow rate did not always increase with the 10dB increment of vocal intensity. He pointed out that the low rate of increase at the low pitch was in contrast with the relatively high rate of increase in the flow rate which was observed generally at the high or the medium pitch.

Scully (1971) designed a study which mainly aimed at relating "aerodynamic features of articulation for /s/ and /z/ to the tense-lax features of articulation and of acoustic output." [p.187]. Oral pressure and volume airflow through the mouth were measured simultaneously for an English speaker. The tense-lax feature was seen to operate 'in the expected way' at the acoustic stage, but it applied 'in the opposite direction' at the articulatory stage. The timing for the movements of the tongue for [s] was almost the same as that for [z] when both segments occurred in similar phonetic contexts. Airflow and pressure were shown to be significantly greater for [s] than for [z]. Still, aerodynamic measures indicating the degree of tongue occlusion showed no significant differences between the two consonants. The acoustic difference between [s] and [z] was quite obvious in that voicing was made to continue through all or nearly all of [z] but it entirely diminished during [s]. In conjunction with the amount of tongue movement and the timing of the tongue movements, both consonants were seen to be similarly articulated in the supraglottal vocal tract.

The results of the oesophageal pressure experiment indicated that the tracheal and oral pressures were approximately equal during the production of [s], and that they did not seem to differ very much for [s] and [z] in corresponding contexts.

Within the framework of a computer model of speech production constructed by Scully (1975) and which was then modified and elaborated (Scully, 1979), it has been demonstrated that "durational differences of acoustic segments comparable with those of real speech arise from voiced-voiceless pairs of laryngeal articulations, as a result of differences at the aerodynamic stage of speech production." [p.233]. Spectrograms and intensity traces that were obtained from the experimental work indicated that acoustic events corresponded to but were not identical with supraglottal articulator occlusions of the vocal tract. This was due to the fact that "all the constrictions in the vocal tract contribute to the aerodynamic conditions at any one point in the vocal tract." [loc.cit].

The model was designed to investigate vowel and consonant durations in sequences containing voiced and voiceless plosives and fricatives. Movements of articulators were represented by a small number of constrictions of the vocal tract.[1] The results obtained from the model showed that durational contrasts similar to those found in real speech arose simply as a consequence of contrasting aerodynamic conditions. Scully (op.cit.) stated that durational differences for vowel-consonant sequences with voiced and voiceless fricatives could be explained on the basis that subglottal and supraglottal articulator actions were the same for both. She added that "the observed aerodynamic differences can arise directly as a result of the single glottal articulatory distinction while the observed acoustic differences arise as a result of the aerodynamic conditions and not directly from the myodynamic articulatory stage of speech production." [p.288].

The function of the velopharyngeal mechanism has been studied by a number of investigators (e.g. Lubker and Moll, 1965; Warren, 1964; Warren and DuBois, 1964; Warren and Devereux, 1966; Smith and Weinberg, 1980) who employed various techniques to infer the cross-section area constriction at the velopharyngeal orifice. For instance, Lubker and Moll (1965) developed a method to measure oral and nasal airflows separately, yet synchronously, in conjunction with cinefluorographic observations. They devised a facemask to separate the oral and nasal airflows and to determine the mask effects on articulatory positioning. Their aim was to determine "a technique whereby such a mask can be used in synchrony with cinefluorographic techniques to investigate the relationships between oral and nasal airflows and various articulatory positions." [p.258].

---

[1] In her study of (1979), Scully stated that "Articulatory movements are represented by changes in cross-section area of a small number of constrictions of the vocal tract." [p.37].

Lubker and Moll (op.cit.) pointed out that the facemask had definite effects on labial and mandibular movement and on maximum rate of syllable production.[1] They commented that it was difficult "to assess the significance of the mask's effect on maximum rate of syllable production, since this rate is not characteristic of normal speech production." [p.268]. No effects were noticeable on tongue movement. They claimed that the advantage of the facemask as a device to describe the airflow characteristics of normal speech might be limited because of the restrictions imposed by the mask on articulatory movements. They rejected the idea of using larger facemasks because of two major problems. First, "by enlarging the mask, a vast amount of dead space is added to the system," and secondly, "any increase in size of the facemask also will increase the amount of space that must be sealed off between the oral and nasal cavities and, therefore, will make it more difficult to achieve an air-tight seal between the cavities." [p.269].

Later, Scully (1969) recognized the same difficulties pertaining to the use of the facemask in experimental work on aerodynamics of speech production. She distinguished two types of problems associated with the use of the aerometer and data interpretation: practical and theoretical problems. Among the most important practical problems that affect the interpretation of airflow traces, she mentioned

---

1. Also see Warren and Wood (1969) who state that "facemasks tend to restrict mandibular movement and this may cause some distortion of speech production." [p.467].

the following points:[1]

1. Adjustment of the facemask to fit differently shaped faces.
2. Uncertainty of the mask's fit being completely airtight during speech.
3. Possibility of the mask restriction to jaw movements.
4. Possibility of introducing errors into airflow traces of speech due to changes in the volume of the mask space when the jaw moves while the mouth is closed.
5. Mask reduction of auditory feedback to the speaker.
6. Difficulty of checking the baseline during speech due to the in and out combination of oral airflow on one channel.

Quite recently, Karnell and Willis (1982a) have reported an investigation to examine vowel context as a variable affecting consonantal peak intraoral air pressure during the production of selected consonants. They used speech samples consisting of consonant/vowel combination in which the consonantal segment was either [p] or [b], and the vowel segment was either [a] or [u]. The instrumentation they employed in their study is known as the 'Agnellograph Pressure Translator'; a device which has been newly developed by the Kay Elemetrics Company and which has been described by the same authors in another paper (Karnell and Willis, 1982b).

Following previous studies (e.g. Warren, 1964; Arkebauer, 1964; Subtelny et al., 1966; Brown and McGlone,

---

[1] For further information see Scully, 1969, pp.58-60.

1969; Bernthal and Beukelman, 1978) they stated that the voiceless bilabial plosive [p] was found to occur with higher peak intraoral air pressure than its voiced counterpart [b]. Both consonants showed higher mean peak intraoral air pressures when in context with [u] than when in context with [a]. Their findings suggested that modifications in the dynamic activity of the vocal tract have an effect on intraoral air pressure, and their study demonstrated that the coarticulatory effects of vowel context on plosive bilabial consonants can be observed and quantified through measurement of peak intraoral air pressure.

6.3 Experimental Procedure

6.3.1 Selection of Stimuli and Recording Technique

The same words referred to in our acoustic experiment (see section 3.2.1) and in our experiments on speech synthesis (see section 4.3.5) have been used in the present investigation. The chosen words represented two oppositions involving different contoid types in which contrasts between single and geminate fricatives and plosives were used. The single consonants [s] and [d] versus their geminate counterparts [ss] and [dd], each used in word-initial and word-medial positions, were investigated. All words selected for this experiment were bisyllabic with primary stress on their first syllables. They were : /sabit/ and /ssabit/, /ħasan/ and /ħassan/, /darub/ and /ddarub/ and /badal/ and /baddal/. Each word was said twice within one and the same constructed carrier sentence, i.e. /'kitbi——— 'sit mar'raat/.

As mentioned earlier, the idea behind using a carrier sentence instead of enunciating words in their isolate forms is because the context of the sentence facilitates closer control of every feature and gives greater consistency and renders the sound segments more natural.[1] This, of course, will help maintain normal stress and intonation patterns. Besides, using these stimulus items in the same carrier sentence will minimize the variability of the data and this, in turn, will help the investigator to keep, to a certain extent, the extraneous variables under his control. The accentual pattern of the sentence was kept to be the same, as far as possible, with the same tempo everytime the sentence was spoken, and a moderate breath was taken between utterances.

As part of the present experiment, a list, making a total of 80 items, was recorded in the recording studio of the Phonetics Laboratory on high quality recording tape (type Ampex 1200). The speed of the tape was 7½"/sec, and the microphone used was of type AKGD 202E1. Gaps were left between recorded items so that this recording procedure could be used as a playback while conducting the experiment later.

The recorded list made five runs. Each run contained 16 items, and in each run the 16 recorded items were randomized anew every time. The dummy carrier sentence /'kitbi ki'taab 'sit mar'raat/ 'write book six times' followed by a silent period of almost the same temporal length as that of the dummy carrier sentence and by the word 'beat' pronounced, rather slowly, three times, was said at the beginning and

---

[1] Cf. Odisho, 1975, p.145.

end of every individual run. The word 'beat' was intentionally inserted after the silent period to help get the baseline, i.e. zero line, for the 'area' traces since there is complete closure of the articulators in a carefully produced plosive consonant, and also to check that the oral and nasal airflow portions of the double mask are well separated. Similarly, the silent period helped to identify the baselines for both oral airflow and oral pressure traces. Both the dummy carrier sentence and the word 'beat' were excluded while making segmentation and measurements at later stages.

6.3.2 Subjects

The author of the present study was his own subject for all studio recordings as well as for this pilot experiment. See section 3.2.2 for further information as to his linguistic background.

6.3.3 Instrumental Set-up

6.3.3.1 Airflow

Airflow traces were obtained by fitting a rubber mask firmly to the subject's face; around the nose and the mouth. The mask was pressed tightly on the subject's face to minimize the escape of air through the soft rubber frame of the mask. Inside the mask there was a relatively solid rubber flange which thoroughly separated the nose from the mouth. On the external surface of the mask two Mercury flowheads (type F1OOL) were affixed; one to the upper part and one to the lower part. The upper flowhead was for nasal airflow and the lower one was for oral airflow. Each of the two flowheads was fitted with a Gaeltec pressure transducer (type 8T, ± 2 cm H_2O).[1] The output signals of the two transducers were fed into a Gaeltec control unit, which acts as an amplifier, and then, inside an aerodynamic speech analyzer, they were low-pass filtered at 50 Hz.

In order to be quite certain that there was no leakage of air through the solid rubber flange separating the nose from the mouth, a preliminary checking was carried out by breathing in and out first through the nose and then through the mouth. There were two airflow traces which were graphically recorded on two separate mingograph channels.

6.3.3.2 Oral Pressure[2]

Oral pressure traces were obtained by using a soft sensing polyethylene tube with one opening near its end (type Franklin's oesophageal tube, size 8ch-257). At first, a small polyethylene tube with a bore of about 1.5 mm was tried. But it was found preferable, for the subject's comfort, to replace the small tube by a slightly larger one, and satisfactory results were obtained. The external diameter of the larger tube was about 3 mm and its internal diameter was about 2.5 mm. In order not to hinder normal articulation, the subject inserted the tube through his nose down to the pharynx rather than through his mouth. He also found it more convenient to insert the tube through his nose because he

---

[1] The whole assembly of the flowhead and the pressure transducer is normally called 'pneumotachograph'.

[2] Henceforth the terms 'oral' and 'intraoral' will be used synonymously.

discovered that he had a highly sensitive velum which was easily irritated at the touch of a plastic tube. The distance between the position of the free end of the tube in the pharynx and the naris was about 15 cm.

Before starting the experiment an initial attempt had been made to insert the polyethylene tube through the nose with the help of a lubricator. The experiment was then carried out by inserting the tube without it. The air compressed behind the vocal tract constriction was transmitted by means of the polyethylene tube to a Gaeltec pressure transducer (type 3CT, ± 10 cm H_2O) through the small opening near the closed end of the tube.

In order to measure the pressure drop across the tongue constriction (ΔP) as pressure difference between the tube behind the constriction and the reference tube in front of the constriction, one end of the reference tube was inserted into the mouth division of the mask; while the other end was linked to the Gaeltec pressure transducer. The output signals of the transducer were fed into a Gaeltec control unit. The signals were low-pass filtered at 50 Hz inside an aerodynamic speech analyzer which was also linked to the mingograph. The resulting traces of the oral pressure were displayed on channel 4 of the mingograph.

6.3.3.3 Area

An aerodynamic method of inferring aspects of the subject's articulatory control was used in this experiment. If the volume flow rate of air through the constriction (U) and the pressure drop across the constriction (ΔP) are known, then the cross-sectional area of constriction (A) may be inferred, approximately, by means of the 'working' orifice equation:

$$U = k.A. \sqrt{\Delta P}$$

where k is an empirical constant (Warren and DuBois, 1964).

In this experiment, the method was used to give, approximately, the timing of the smallest constriction area formed by the tongue and the denti-alveolar region of the palate in the production of [s] and [ss]. The magnitude of the constriction at its smallest was compared for [s] and [ss], also.

Oral 'area' traces were obtained from an aerodynamic speech analyzer. It is an analog computer device designed and constructed by R.W. Caley, Electronic Instrument Design. This device accepts, as input, electrical signals proportional to the volume airflow rate (U) and the pressure drop (ΔP). It gives, as output, an electrical signal proportional to $U/\sqrt{\Delta P}$. All the signals inside the aerodynamic speech analyzer were low-pass filtered at 12 dB/octave with a cut-off frequency at 50 Hz. The actual magnitude of constriction area (A) in cm^2 is given by the following formula:

$$A(cm^2) = \frac{kU(cm^3/sec)}{\sqrt{\Delta P(cm\ H_2O)}}$$

In this study, we used van den Berg et al.'s (1957) value for k (also used by Stevens, 1971). Therefore, an estimate for actual cross-sectional area of a constriction is obtained by using the formula below:

$$A(\text{cm}^2) = \frac{0.00076 \times U}{\sqrt{\Delta P}}$$

where U is in cm^3/sec and ΔP is in cm H_2O.

The 'area' trace was used only to locate the minimum value of constriction area, replacing the laborious sampling and calculation otherwise required. At the times of minimum area thus located, the relative magnitudes of constriction area were calculated from the calibrated traces of airflow and pressure drop. It is possible to employ the aerodynamic speech analyzer with the airflow and pressure transducer systems because each of these has a linear calibration. The 'area' traces were displayed on channel 5 of the mingograph.

6.3.3.4 Electro-glottograph

Mingograph traces for the electro-glottograph were obtained by placing two electrodes (F-J Electronics) horizontally on both sides of the external surface of the subject's larynx. The two electrodes were held firmly in place by a velcro strap which was tied around the neck of the subject. The vibrations of the vocal folds were transmitted into electric signals via the two electrodes to an electro-glottograph (F-J Electronics). The input to the electrode voltage was set at 4 for 'coarse' and at 6½ for 'fine'; the output associated with the vocal folds was set at 5½.

In order to get an audio version of the behaviour of the vocal folds, the electric signals were transmitted from the electro-glottograph to track 2 of the Revox A77 tape-recorder. The input to track 2 was set at 5½ Aux, and the output was set at 9¼. The balance of right/left was set at 12 o'clock. The output signals were also fed from the electro-glottograph into the mingograph which graphically recorded electro-glottograph traces on channel 6.

6.3.3.5 Duplex Oscillogram

Duplex oscillogram traces were obtained by placing a microphone (type Electret) outside the mask in front of the subject's face. In order to have a recorded version of the stimuli presented in this experiment and to discard any wrong or ambiguous items before performing our segmentation and measurements, the microphone was linked to the tape-recorder and all the acoustic data of the experiment were recorded on high quality tape (type Ampex 1200) by using track 1 of the Revox A77 tape-recorder. The input to track 1 was set at 4½, and the output was set at 9¼. From the tape-recorder the signals were transmitted to a fundamental frequency, i.e. FO, meter (F-J Electronics) which, in turn, sent its signals to the mingograph. The balance control knob for the duplex oscillogram on the FO meter was set at 10 o'clock. Duplex oscillogram traces were displayed on channel 1 of the mingograph. Fig. 6.1 illustrates the whole instrumentation employed in this experiment.

6.4 Segmentation and Measurements Criteria

6.4.1 Introduction

For the sake of getting consistent results in our

Fig. (6.1) **Block diagram of instrumental set-up used in the pilot experiment to investigate the aerodynamics of single and geminate consonants**

Fig. (6.1)

Description:

A	Flowhead for nose airflow (Mercury).
A'	Flowhead for mouth airflow (Mercury).
B	Pressure transducer for nose flowhead (Gaeltec).
B'	Pressure transducer for mouth flowhead (Gaeltec).
C	Pressure transducer for pressure tubes (Gaeltec).
D	Electro-glottograph (F-J Electronics).
E	Microphone placed outside the facemask (type : Electret).
F	Control box and amplifier for airflow (Gaeltec).
G	Control box and amplifier for oral pressure (Gaeltec).
H	2-track tape recorder (Revox A77).
I	FO (Transpitch) meter to give duplex oscillogram (F-J Electronics).
J	Aerodynamic speech analyzer with low-pass filtering (Richard Caley).
K	Mingograph (Mingograf No. 803/Siemens Elema).

Mingograph traces:-

Ch.0	Timer.
Ch.1	Duplex oscillogram (gain set at 7¼).
Ch.2	Volume flow rate of air through the nose (gain set at 9)
Ch.3	Volume flow rate of air through the mouth (gain set at 6½).
Ch.4	Oral pressure with air in front of the mouth as reference (gain set at 6).
Ch.5	'Area' of oral constriction, derived from oral airflow and oral pressure (gain set at 3½).
Ch.6	Glottograph signal (gain set at 9).

experiments on the aerodynamic and articulatory activities of the vocal tract during speech, we present certain formulae for segmenting and measuring those items under investigation. Aerodynamic and articulatory segmentation and measurements were based on oral airflow, oral pressure and oral 'area' traces. Points for 2 articulatory events were established on these traces; these were taken to represent the beginning and end of an articulatory segment. In this way the articulatory durations of the segments and utterances to be analysed were established. By reference to these points, both aerodynamic and articulatory measurements were made.

All oral airflow, oral pressure and oral 'area' traces have a linear calibration. That is to say, when the trace looks twice as high on the mingograph paper, the actual physical thing being recorded really is twice as large (Scully, 1982). For this reason, one can measure airflow, oral pressure and oral area (in mm) up from the baseline, i.e. zero line, to get the contrasting patterns.

For the purpose of presenting our tables, all the numbers indicating measurements of duration, oral pressure and oral airflow have been rounded up to the first decimal place. This is because we first made our measurements in millimeters and then they were converted into milliseconds for duration, cm H_2O for oral pressure and cm^3/sec for oral airflow.

Probably the estimated accuracy of measurements will be as follows:

1. for duration is ± 4 msec;
2. for oral pressure is ± 0.2 cm H_2O; and
3. for oral airflow is ± 20 cm^3/sec.

2 and 3 imply that the estimated accuracy for actual magnitude of constriction area (A) is about ± 0.005 cm^2. This needs to be born in mind when considering the patterns shown in the numbers presented in the tables.

Traces of duplex oscillograms and electro-glottograms have also contributed to our segmentation procedures. They have been of major help in inferring the activities of the vocal folds as well as identifying different phonetic contrasts, i.e. distinctions between voiced versus voiceless segments.

6.4.2 Articulatory Duration of Supraglottal Constriction

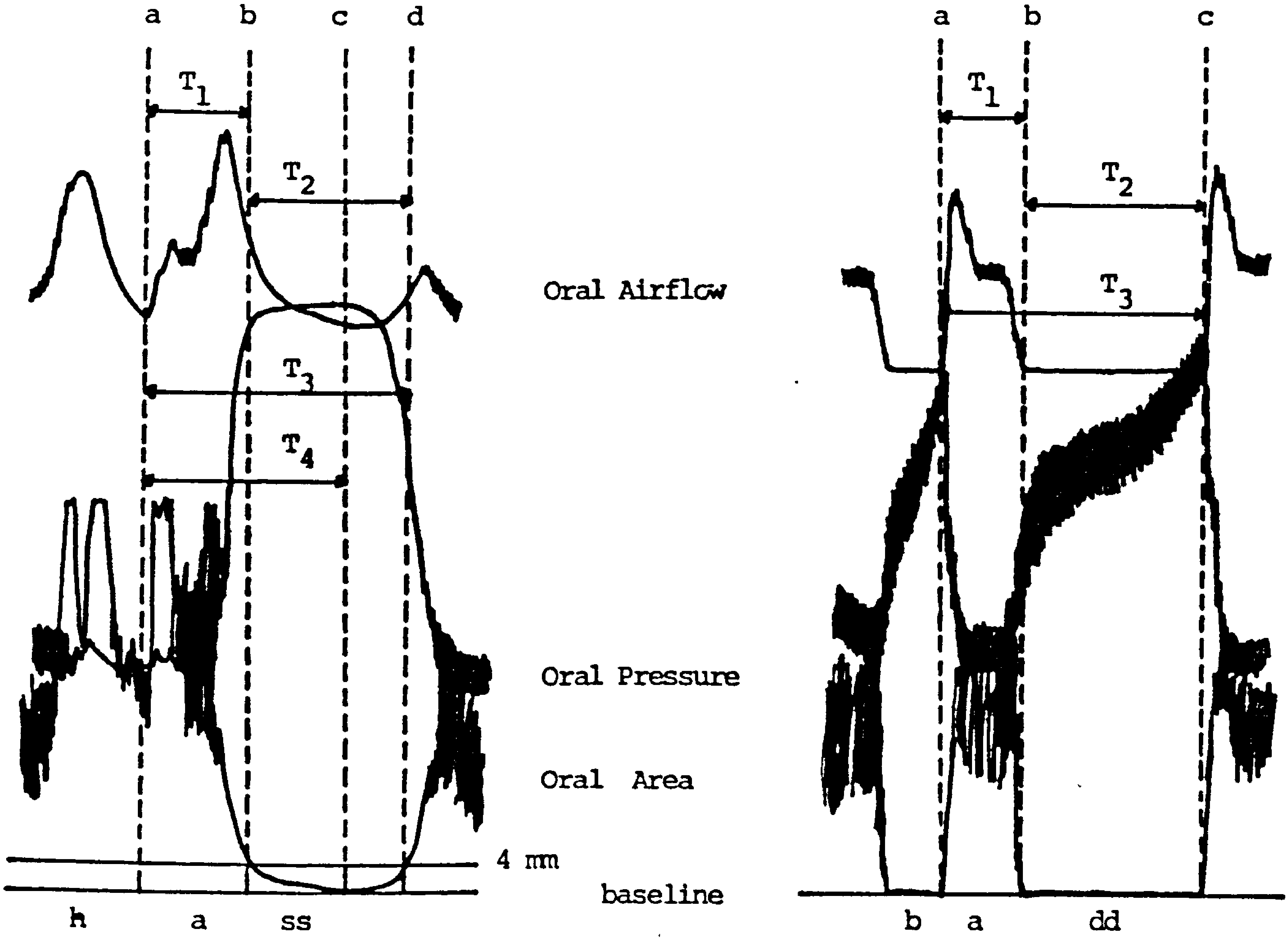
In order to measure the articulatory duration (T_2) of the segments under investigation, a baseline for oral area was drawn. Similar baselines for oral airflow and oral pressure were also drawn. Points were established on the oral 'area' trace to represent the initiation and termination of the articulatory events of [s] and [ss] on the one hand, and of [d] and [dd] on the other in word-initial and word-medial positions. For the two plosives [d] and [dd], the point of initiation is taken where the falling 'area' traces touch the baseline. This fall is taken to denote the end of the preceding vowel and the beginning of the following consonant. At the same time, the termination point of [d] and [dd] is taken as a sudden rise where the 'area' trace departs from the baseline. This rise is taken to denote the end of [d] and [dd] and the beginning of the following vowel.

For the two fricatives [s] and [ss], an arbitrary threshold value was established at 4 mm above the lowest point in the dip of the 'area' trace. The dip of the 'area' trace is taken to represent the smallest oral constriction area value for a fricative. After locating the smallest oral area value, a synchronizing line was drawn up the mingograph squared paper so as to measure airflow and pressure (in mm) at this point of time up from zero line, i.e. the baseline. The smallest oral area value was also measured at this point (in mm) up from zero line. Measurements of duration were made (in msec) between points in time where the 'area' trace crosses the chosen arbitrary threshold value. These points were taken to represent the articulatory duration of supraglottal constriction for [s] and [ss].

It is worth stating that similar, but not identical, contrasts for the oral 'area' trace for [d] and [dd], and for [s] and [ss] were also noticeable on the oral airflow trace, see section 6.4.3. Figs. 6.2 and 6.3 show how segmentation and measurements of articulatory duration were made on the 'area' trace for the consonants under analysis.

6.4.3 Articulatory Timing

To begin with, by 'articulatory timing' of a word, or a syllable, or a sequence, we refer to the duration existing



Figs. (6.2 and 6.3)

Mingograms of the sequences /ħass-/ in /ħassan/ and /badd-/ in /baddal/ showing the criteria for segmentation and measurements of articulatory timings for the investigated stimuli.

between the articulatory events within and preceding the investigated consonantal segments in our selected stimuli. Therefore, only articulatory events preceding the examined consonants in word-medial position have been studied. For example, the time between the moment of release, i.e. occlusion offset, of an initial plosive and that of a medial one, as in /badal/ and /baddal/, or the time between minimum cross-sectional area (Ac min. for short) of an initial fricative and that of a medial one, as in /ħasan/ and /ħassan/ was measured, see start of segment T_1 in Figs. 6.2 and 6.3.

The objective behind measuring the articulatory timings, i.e. T_1, T_3 and T_4, is to see whether there is a significant difference in articulatory duration between a vowel preceding a single consonant and that of a similar one before a geminate consonant provided that the two consonants should be of the same phonetic manner and place categories.

For [s] and [ss] occurring word-medially, the lowest point on the dip of the 'area' trace, representing Ac min., was taken as the point in time for carrying out our measurements; whereas for the initial voiceless pharyngeal fricative [ħ] the time of the minimum value of Ac was based on measurements made by using the dip of the oral airflow, i.e. minimum value of volume airflow rate (Uo min. for short), since it was found that the point in time at Ac min. very often coincided with a corresponding point of Uo min. It was not possible to obtain oral pressure traces for [ħ] so that the 'area' trace could not be used here.

In Fig. 6.2, a b (T_1) represents the articulatory timing (in msec) for vowel [a] preceding [s] and [ss]; a d (T_3) represents the articulatory timing for the sequences /ħas-/ in /ħasan/ and /ħass-/ in /ħassan/; a c (T_4) represents the articulatory timing between Ac min. of [ħ] and that of [s] and [ss]; b d (T_2) represents the articulatory duration of [s] and [ss], as discussed in section 6.4.2.

In the case of [d] and [dd], three points in time have been considered, see Fig. 6.3 . Point a represents the time of the occlusion offset, i.e. the release, of word-initial [b] as established on the 'area' trace where it suddenly departs from the baseline upwards; point b represents the time of the occlusion onset of word-medial [d] and [dd] where the 'area' trace suddenly goes down to reach the baseline; point c, on the other hand, represents the time of the occlusion offset of the two consonants [d] and [dd] where the 'area' trace moves sharply from the baseline upwards. Again, it was obvious that both downward and upward movements of the 'area' trace, indicating the initiation and termination of [d] and [dd], coincided with similar movements for the two plosives on the oral airflow traces. It was also observed that the start of the sudden rise of the 'area' trace matched closely movement of the peak pressure for the two plosives on the oral pressure traces.

As for word-initial [b] in /badal/ and /baddal/, the sudden upward movement of the 'area' trace, indicating the offset of the plosive occlusion, was taken to represent the starting point in time for articulatory timing measurements.

In cases where the beginning of the sudden rise of the 'area' trace was not clearly identifiable, the peak pressure on the oral pressure trace was regarded as the key to indicate the starting point instead. In Fig. 6.3, a b (T_1) represents the articulatory timing of vowel [a] preceding [d] and [dd]; a c (T_3) represents the articulatory timing between occlusion offset of [b] and that of [d] or [dd], i.e. the articulatory timing for the sequences /bad-/ in /badal/ and /badd-/ in /baddal/; b c (T_2) represents the articulatory duration of [d] and [dd], as explained in section 6.4.2.

6.4.4 Oral Pressure

In order to find out the pressure values at peak oral pressure points, a baseline for oral pressure traces was drawn for all items to be analyzed, namely [s] versus [ss] and [d] versus [dd], both in word-initial and word-medial positions. A vertical line was drawn between the baseline and the peak points of the oral pressure traces for all segments under investigation. The distance from the baseline up to the peak point was measured (in mm). Measurements (in cm H_2O) were based on a linear calibration of oral pressure traces.

6.4.5 Minimum Cross-Sectional Area of Constriction

For the purpose of estimating the timing of the minimum area of constriction (i.e. Ac min.) created by the articulatory organs while producing the fricative consonants, the dip of the 'area' trace was taken to represent the moment of the smallest value of constriction area for fricatives. The actual magnitudes of constriction area were calculated at the points of time of minimum area which were extracted from the calibrated traces of both oral airflow and oral pressure.

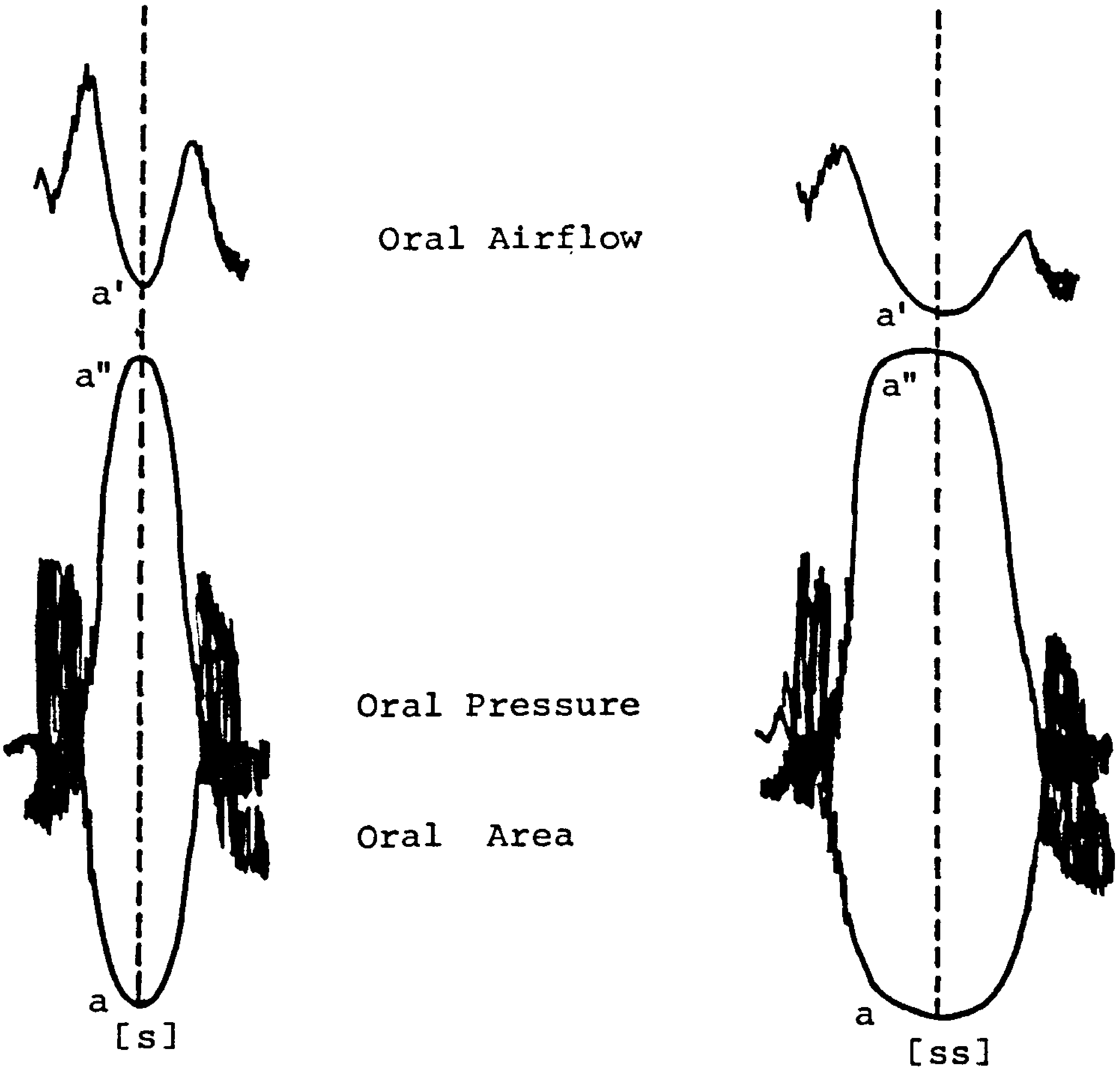
In Figs. 6.4 and 6.5, for instance, point a represents the time of the smallest value of cross-sectional area of constriction (i.e. Ac min.) for the single fricative [s] and its geminate cognate [ss]. a' and a" are the corresponding points of time on the oral airflow and oral pressure traces which represent, respectively, the value of volume airflow rate (Uo for short) and the value of oral air pressure (Po for short) relative to the air pressure just in front of the mouth at the time of a. Following previous works of some researchers (for example Warren and DuBois, 1964, p.52; Lubker, 1969, pp.218-219; Scully, 1971, p.189; Smith and Weinberg, 1980, p.277), the cross-sectional area of constriction can be estimated by means of the following general orifice equation:[1]

$$A = \frac{kU}{\sqrt{\Delta P}}$$

where A is constriction area, k is an empirical constant, U is the volume flow rate of air, and ΔP is the pressure drop across the constriction. Therefore, the relative magnitude of A_c at point a has been calculated by the value of U at point a' divided by the value of $\sqrt{\Delta P}$ at point a".

---

[1] More details about this equation have been mentioned in section 6.3.3.3.



Figs. (6.4 and 6.5)

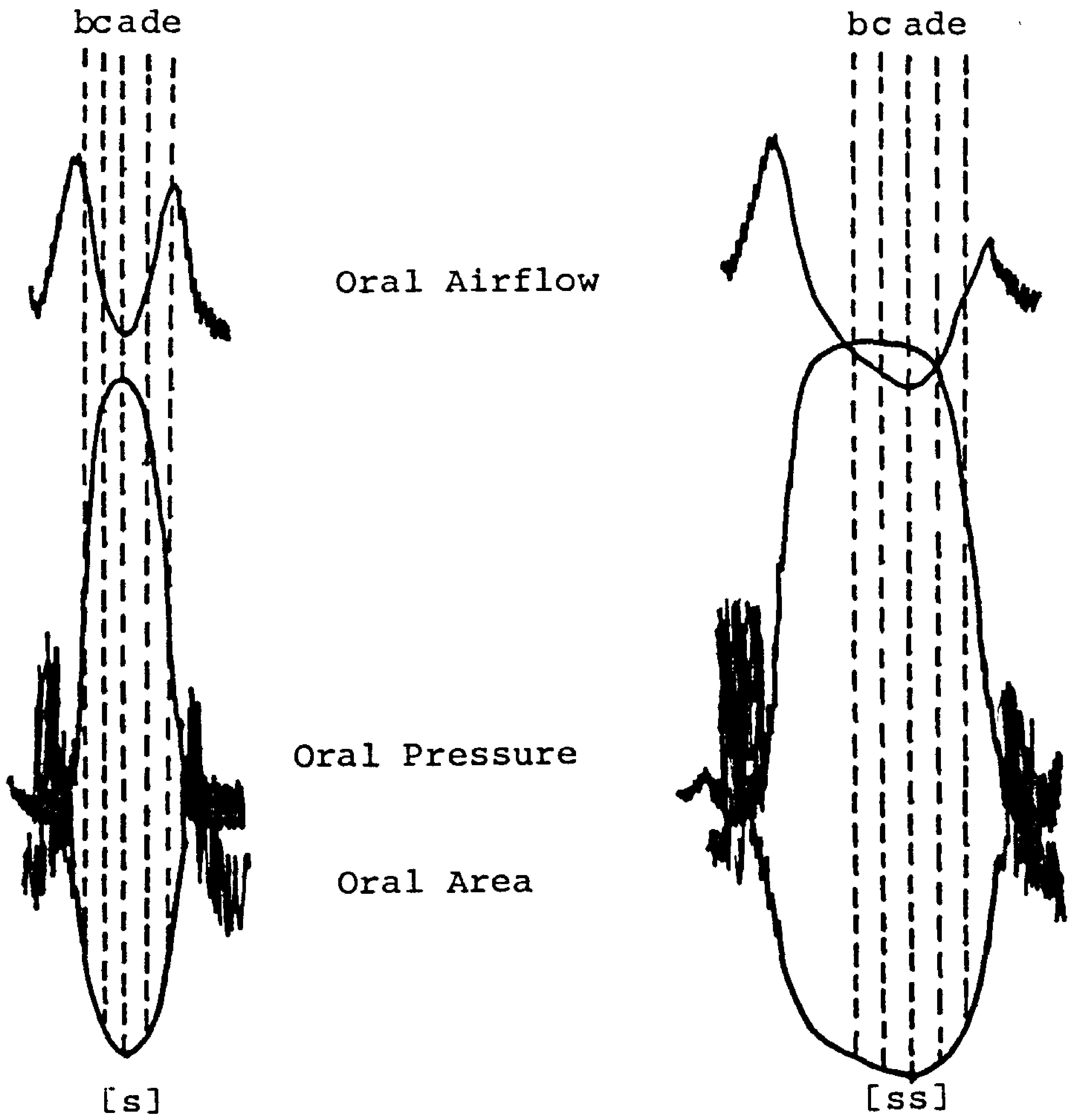
Mingograms of [s] in /ħasan/ and [ss] in /ħassan/ showing the criteria for segmentation and measurements of the minimum cross-sectional area of constriction (Ac) estimated from oral airflow and oral pressure measurements. Point a corresponds in time to Ac min., point a' corresponds to the value of volume airflow rate (Uo) at the time of point a, point a" corresponds to the value of oral pressure (P_o) at that point.

To give approximately the timing of the smallest constriction area created by the supraglottal articulators in the production of fricative consonants and to confirm the accuracy of our choice of the dip of the 'area' trace as the minimum value of $U/\sqrt{\Delta P}$, we examined the plausibility of that choice by following a procedure where a number of points were chosen randomly on the 'area' trace, in addition to point a, near the dip of the trace for single [s] in /ḥasan/ and its geminate counterpart [ss] in /ḥassan/, as shown in Figs. 6.6 and 6.7. Both Uo and Po were measured at these various points of time. Values of constriction area were, then, calculated at these points, and it was found that point a gave the smallest value of cross-sectional area of constriction in the two examples under consideration. See tables 6.2 and 6.3.

While performing our segmentation and measurements, we observed that point a at Ac min. always coincided with point a' of Uo min, i.e. the dip of the oral airflow trace. We, therefore, took advantage of this coincidence between the two points and considered the dip of the oral airflow trace to represent Ac min. for back consonants, such as the voiceless pharyngeal fricative [ḥ] in /ḥasan/ and /ḥassan/, for which Po traces were dififcult to obtain and, thus, the measurements for finding the time of the minimum value of Ac could not be estimated.[1] In section 6.4.3, it has been shown that this representation of Ac min. with the help of Uo min. was of benefit to us in measuring articulatory events for back consonants.

---

[1] This observation has also been reported by Hassan, op.cit. pp.279-282.



Figs. (6.6 and 6.7)

Mingograms of [s] in /ħasan/ and [ss] in /ħassan/ showing how the Ac min. value was estimated. Point a corresponds in time to the minimum of the 'area' trace. Points b, c, d, e were arbitrarily chosen on the 'area' trace. Uo and P_O values were measured at their corresponding points on the oral airflow and oral pressure traces. It was found that point a had the minumum value of Ac $= \frac{k\,Uo}{\sqrt{Po}}$ for all the selected checks made.

Table (6.2) Calculations of Ac at the points in time of a, b, c, d, e on the 'area' trace for the single fricative [s] in /ħasan/ as shown in Fig. 6.6. The results indicate that point a has the minimum value of Ac in all cases.

Points	Uo(cm^3/sec)	Po (cm H_2O)	$Ac = \frac{kUo}{\sqrt{Po}}$ (cm^2)
a	454.5	14.4	0.091
b	666.7	11.9	0.147
c	515.2	13.3	0.107
d	575.8	12.5	0.124
e	727.3	10.8	0.168

Table (6.3) Calculations of Ac at the points in time of a, b, c, d, e on the 'area' trace for the geminate fricative [ss] in /ħassan/ as shown in Fig.6.7. The results indicate that point a has the minimum value of Ac in all cases

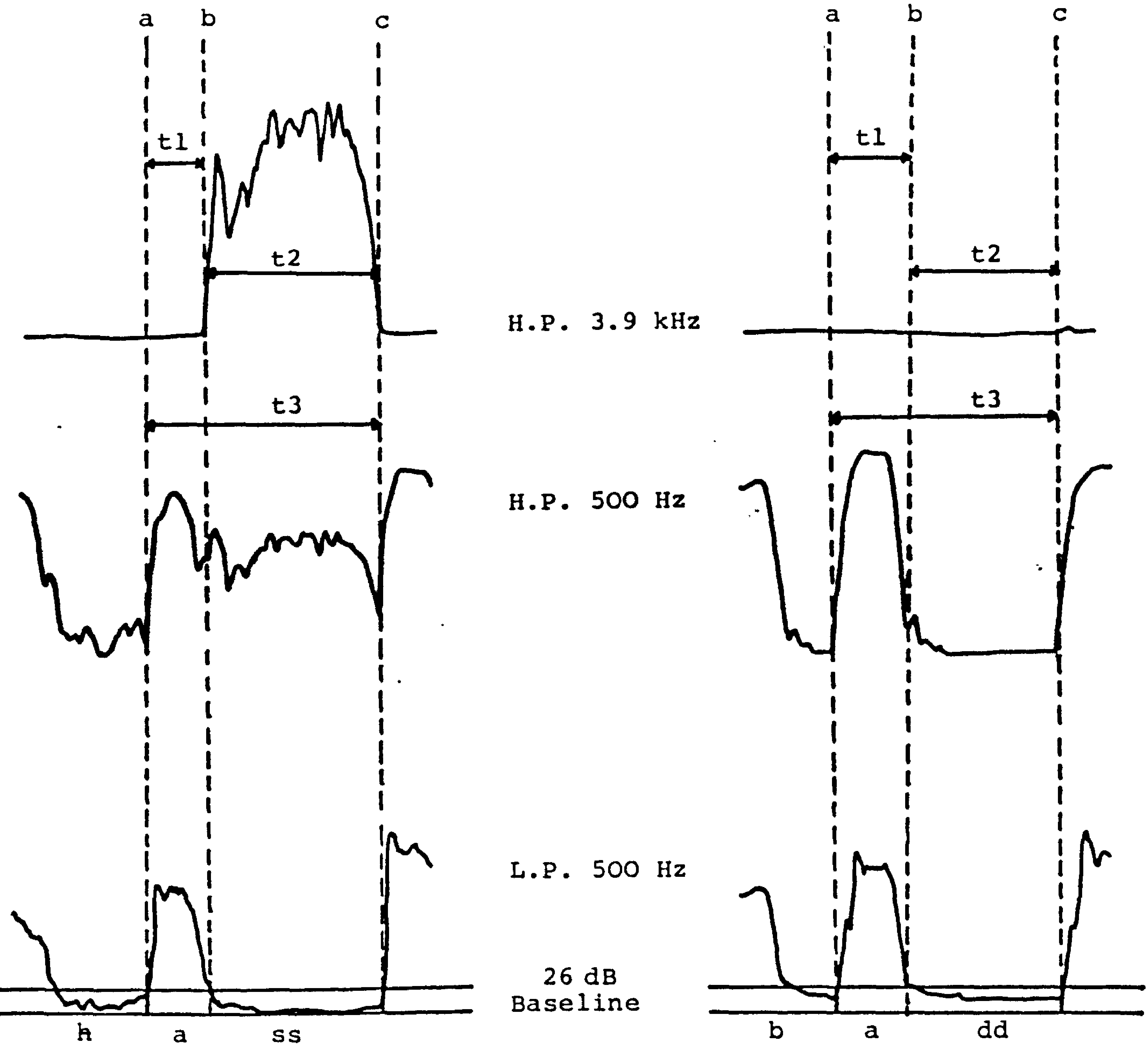
Points	Uo(cm^3/sec)	Po (cm H_2O)	$Ac = \frac{kUo}{\sqrt{Po}}$ (cm^2)
a	272.7	15.5	0.053
b	393.9	15.3	0.077
c	303.0	15.6	0.058
d	333.3	14.7	0.066
e	454.5	12.2	0.099

6.5 Acoustic Measurements

6.5.1 Instrumental Set-up and Segmentation Criteria

The quantitative acoustic data for the stimulus items of this experiment were separately obtained at a later stage by using the same instrumentation, as described in section 3.2.3. The intensity meter was employed, and three types of intensity level traces were graphically recorded via the mingograph. Channel 1, of the mingograph, displayed intensity level traces high pass filtered at 3.9 kHz (log scale), channel 2 displayed intensity level traces high pass filtered at 500 Hz (log scale), and channel 3 displayed intensity level traces low pass filtered at 500 Hz (linear scale).

The segmentation and measurements procedures used in the present investigation were the same as those discussed in our acoustic experiment of Chapter Three. Briefly, the acoustic durations of the investigated consonants and of their preceding vowels were made (in msec) on the low pass intensity level traces filtered at 500 Hz with an arbitrary threshold value set at 26 dB above the baseline. The high pass intensity level traces of both types, i.e. those filtered at 500 Hz as well as those filtered at 3.9 kHz, were of great support, especially when determining the initiation and termination points of fricative consonants. It was not necessary to establish an arbitrary threshold value for the H.P. 3.9 kHz trace since it rose and fell sharply from and to its baseline. Full explanation of how our acoustic segmentation and measurements criteria were



Figs. (6.8 and 6.9)

Minograms (intensity level traces) of the sequences /ħass-/ in /ħassan/ and /badd-/ in /baddal/ showing the criteria for segmentation and measurements of acoustic timings and durations for the investigated stimuli.

carried out can be referred to in section 3.2.4.

Similar to the terminology used in describing our articulatory measurements criteria, Figs. 6.8 and 6.9 show that a b (t_1) represents the acoustic timing (in msec) for vowel [a] preceding [s] and [ss], or [d] and [dd]; a c (t_3) represents the acoustic timing for the sequences /ħas-/ in /ħasan/ and /ħass-/ in /ħassan/, or /bad-/ in /badal/ and /badd-/ in /baddal/; b c (t_2), on the other hand, represents the acoustic duration of [s] and [ss], or [d] and [dd]. From this description, then, it is apparent that we distinguish between 'acoustic duration' on the one side, and 'acoustic timing' on the other exactly in the same way as we distinguish between 'articulatory duration' and 'articulatory timing', see sections 6.4.2 and 6.4.3.

6.6 Statistical Test Used

The results of this pilot experiment were subjected to the same statistical treatment of the Mann-Whitney U-test used in our experiments on speech synthesis. The two levels of significance for rejecting H_0 reported in those experiments have also been used in our present investigation. Therefore, differences between single and geminate consonants at $p \leq 0.01$ will be described as statistically highly significant; whereas differences at $p > 0.01$ and $p \leq 0.05$ will be described as statistically significant. At the same time, differences at $p > 0.05$ will be described as statistically non-significant. The first and the second levels of significance will be similarly marked by the signs (**) and (*), respectively. The non-significant level will be left unmarked. See section 4.4 for full details of the statistical test used in this experiment.

Generally, statistical tests are tools used to analyze results that may lead to inferences concerning the reliability of the difference between groups of scores. Statisticians (e.g. Miller, 1975, p.119) state that there is always a grave danger that statistical significance may blind the experimenter to possible weaknesses of design and conceptualization which underly the numerical results. They also emphasize the fact that the practical and theoretical significance of the results is determined by many considerations which are external to the statistical analysis. Among these considerations are the adequacy of the design, the magnitude of the experimental effect and the nature of the subject population. Therefore, in examining the results of our pilot and main experiments we do not only describe the p values in the tables, but we also consider the overall pattern trends, i.e. the actual measurements of the two sets (taking the accuracy of measurements into account), their mean values and their standard deviations, before we draw our conclusions.

6.7 Analysis of Results

6.7.1 Oral Pressure

Tables 6.4 - 6.7 show the measurement values of peak oral pressure. Each table contains the ten tokens that have been analyzed, the mean values of their peak oral pressure, the standard deviation, the U value of the statistical treatment of the Mann-Whitney U-test, and the p value indicating the significance of the differences between the stated measurement values.

Table 6.4 shows that there are significant differences in peak oral pressure between the single fricative [s] and its geminate cognate [ss] when both are in word-initial position, where p value is at 0.0170. Table 6.5, on the other hand, shows the differences in the values of peak oral pressure between the two segments occurring word-medially. In contrast to the results indicated in table 6.4, table 6.5 shows that these differences are statistically non-significant with p value at 0.1822.

Table 6.6 indicates considerable differences in peak oral pressure between the single plosive [d] and its geminate partner [dd], when the two segments occur word-initially. The differences are statistically highly significant, where p value is at 0.0001. Table 6.7 indicates that the differences in peak oral pressure are still statistically highly significant when both consonants occur word-medially, with p value at 0.0002.

Generally, the statistical evaluations indicated in the results of the four tables show clearly that peak oral pressure is greater for a plosive consonant than for a fricative consonant in comparable context.

6.7.2 Oral Airflow and Oral Pressure at Ac minimum

Tables 6.8 - 6.11 show the results for aerodynamic conditions, i.e. volume airflow rate and oral pressure at the minimum point of cross-sectional area associated with the production of the single fricative [s] and its geminate counterpart [ss] used in word-initial and word-medial positions. Tables 6.8 and 6.9 illustrate that differences in the volume of airflow rate (Uo) between [s] and [ss], used either word-initially or word-medially, are statistically highly significant, where p values are at 0.0044 and 0.0004, respectively.

For oral pressure (Po), table 6.10 illustrates that differences between word-initial [s] and word-initial [ss] are statistically significant, with p value at 0.0204. Table 6.11, on the other hand, shows that the differences in the amount of oral pressure at the minimum point of cross-sectional area between word-medial [s] and word-medial [ss] are statistically non-significant, with p value at 0.1582. It is worth noting that tables 6.10 and 6.11 demonstrate similar trends to those presented in tables 6.4 and 6.5, where differences in oral pressure are statistically more significant as between single and geminate fricatives in word-initial position than as between those used word-medially. These similarities are not surprising, since the pressures have been measured very close together.

However, an inspection of tables 6.12 and 6.13 shows that differences in Uo at Ac min. i.e. Uo min., between [s] used word-initially and that used word-medially are statistically non-significant, with p value at 0.1328; whereas for the geminate [ss], the differences are shown to be significant, where p value is at 0.0240. At the same time, tables 6.14 and 6.15 show that differences in Po at Ac min., i.e. Po min., between [s] occurring in word-initial position and that occurring in word-medial position are also statistically non-significant. The same fact applies to [ss], except that in the case of [s] the differences are entirely negligible, where p value is at 0.7019.

6.7.3 Minimum Cross-Sectional Area of Constriction

Tables 6.16 and 6.17 show significant differences in Ac minimum between the single fricative [s] and its geminate cognate [ss]. More specifically, the smallest value for cross-sectional area of constriction is considerably larger for [s] than for [ss] when both segments occur either word-initially as in /sabit/ versus /ssabit/, or word-medially as in /ħasan/ versus /ħassan/, where p values are at 0.0081 and 0.0006, respectively. In other words, these results indicate strongly that there is a closer approximation between the tip or blade of the tongue and the hard palate when pronouncing a geminate consonant than when pronouncing a non-geminate consonant. This is, in the ordinary course of events, the reason why there is a greater amount of oral pressure and a lesser volume of airflow rate in the production of a geminate consonant; whereas a non-geminate consonant is accompanied by a lesser amount of oral pressure and a greater volume of airflow rate, see tables 6.4 - 6.11.

It is of interest to remark here that the results presented in tables 6.16 and 6.17 are compatible with those reported in our experiments on palatography. They, therefore, give further support to our palatographic findings that geminate consonants are characterized by having a more extensive and clearer wipe-off, thus implying firmer contact between the articulatory organs, than their non-geminate partners, see section 5.4.

Tables 6.18 and 6.19 show that there are no significant differences in the Ac minimum of [s] and [ss] pronounced word-initially and those pronounced word-medially.

6.7.4 Articulatory Duration of the Supraglottal Constriction

In so far as the articulatory duration of the supraglottal constriction, i.e. T_2, for single fricatives and plosives versus their geminate counterparts is concerned, table 6.20 illustrates significant differences in articulatory duration of the supraglottal constriction between [s] and [ss] when both segments are used word-initially. These differences are also statistically highly significant, with p value at 0.0001, when the two segments occur word-medially, as shown in table 6.21.

Tables 6.22 and 6.23 show that the differences in articulatory duration of the supraglottal constriction for the single plosive [d] and its geminate cognate [dd] have the same trend as that of [s] and [ss]. That is, the differences are statistically highly significant when the two consonants, namely [d] and [dd], are pronounced in word-initial and word-medial positions, where p value is at 0.0001 in both tables.

Considering the results shown in tables 6.20 - 6.23, one may, once again, unhesitatingly claim that they provide a further concrete evidence to the facts presented in our direct palatography experiment, where it was suggested that the more extensive and the clearer wipe-off associated with the production of geminate consonants signifies firmer contact between the articulatory organs in contrast with the less clear wipe-off characterizing the non-geminate consonants. Articulatorily speaking, the results presented in tables 6.20 - 6.23 show very clearly that geminate consonants are considerably longer than their non-geminate partners, no matter whether they are fricatives or plosives, used word-initially or word-medially. Consequently, one can say that the wider and clearer the wipe-off of a segment is, the longer articulation it has.

6.7.5 Articulatory Timing

Tables 6.24, 6.25 and 6.26 show significant differences in articulatory timing between the two segments /ħas-/ in /ħasan/ and /ħass-/ in /ħassan/. In table 6.24, for example, it is shown that differences in articulatory timing (T_1) between Ac min. of word-initial [ħ] and the beginning of word-medial [s] or [ss] are statistically highly significant, with p value at 0.0004. Table 6.26 shows significant differences in articulatory timing (T_4) between Ac min. of word-initial [ħ] and that of word-medial [s] or [ss].

From the articulatory point of view, the results of this table demonstrate clearly that the geminate [ss] is considerably longer than its non-geminate counterpart [s], and that the execution of Ac min. for the geminate consonant remains for a longer period of time than that for the non-geminate consonant. Table 6.25, on the other hand, shows that there are significant differences between T_3, i.e. the articulatory timing between Ac min. of word-initial [ħ] and the end of word-medial [s] or [ss], in the sequence /ħass-/ and that in /ħas-/. It is considerably longer in the former sequence than in the latter one, suggesting that [ss] has a significantly longer articulation than [s]. T_3 equals $T_1 + T_2$, i.e. the articulatory timing of the preceding vowel plus the articulatory duration of the consonant under investigation, respectively.

For the two plosives [d] and [dd], table 6.27 illustrates that there are non-significant differences in articulatory timing (T_1) between the occlusion offset of word-initial [b] and the occlusion onset of word-medial [d] in /badal/ and [dd] in /baddal/. The differences are shown to be completely negligible, where p value is at 0.6613. These articulatory results are highly consistent with our acoustic results reported in our experiment of Chapter Three, see tables 3.21 - 3.24. They are also in agreement with our acoustic results presented in table 6.34 of this chapter. These findings strongly suggest that, on the articulatory level, there is no evidence of temporal compensation occurring between a geminate consonant and a preceding vowel; a fact that has been acoustically discussed in Chapter Three of this study, see section 3.3. They also suggest that a geminate plosive is not necessarily accompanied by an earlier closure phase than that for a non-geminate cognate. Indeed, for the fricatives the closure for the geminate appears to start a little later than for the single fricative. This indicates that our present results are in direct discrepancy with the contentions of a number of researchers examining other languages (e.g. Nooteboom and Slis, 1972; Thananjayarajansingham, 1976). On I.C. Arabic, Hassan (1981) found small but significant differences before single consonants than those before geminate consonants. They are, on the other hand, in agreement with the findings reported by Delattre (1971) and McKay (1980) in that durational differences between vowels preceding geminate consonants and those preceding non-geminate consonants are negligible, see section 3.4.

In the case of T_3, i.e. the articulatory timing between occlusion offset of word-initial [b] and that of word-medial [d] and [dd], table 6.28 shows the closure duration for the geminate [dd] is almost exactly twice that for the single [d]. The differences are shown to be statistically highly significant, with p value at 0.0001. The results of this table give further support to those presented in table 6.23 in that a word-medial geminate plosive has considerably longer articulation than its non-geminate counterpart.

However, an examination of the above-mentioned tables will reveal the fact that T_1 has been found to be considerably longer before [ss] than before [s], suggesting that the articulatory timing of vowel [a] is longer before a geminate fricative than that before a single one; whereas with the plositves, T_1 shows no significant differences before either consonant. The difference between fricatives and plosives in articulatory timing is puzzling, and calls for further investigation.

6.7.6 Acoustic Duration

In tables 6.29 - 6.32 the acoustic durations (t_2) of consonants [s] versus [ss] and [d] versus [dd] are presented. These tables show that the geminate consonants are considerably longer than their non-geminate partners, whether they occur word-initially or word-medially. Tables 6.29 and 6.30, for instance, show that the geminate fricative [ss] is twice as long as its single cognate [s]. The durational differences between the two consonants are found to be statistically highly significant, with p value at 0.0002 both in word-initial and word-medial positions. Tables 6.31 and 6.32 show that the geminate plosive [dd] is three times longer than its single counterpart [d] in identical word positions. The differences in duration between the two consonants are also found to be statistically highly significant, with p value at 0.0001.

The results presented in these tables show the same trends as those indicated in tables 6.20 - 6.23 in that geminate consonants have considerably longer articulations than single consonants. They are also highly consistent with the results found in our acoustic experiment reported in Chapter Three, see section 3.3.

6.7.7 Acoustic Timing

Tables 6.33 and 6.34 show the results of acoustic timing (t_1) between the termination of word-initial [ħ] and [b] and the initiation of word-medial [s] and [ss], or [d] and [dd]. Again, these results apparently have similar trends to those obtained in our acoustic experiment of Chapter Three. That is, the durational differences of vowel [a] preceding [s] and [ss] or [d] and [dd], occurring intervocalically, were found to be negligible. These differences in vowel duration were statistically shown to be non-significant, with p value at 0.3174 in the first table and at 0.3461 in the second table. This is, in fact, more evidence indicating the non-existence of temporal compensation between a geminate consonant and a preceding vowel as has been previously discussed.

In tables 6.35 and 6.36, however, the acoustic timing (t_3) is shown to be considerably longer for the sequences /-ass-/ in /ħassan/ and /-add-/ in /baddal/ than for /-as-/ and /-ad-/ in /ħasan/ and /badal/, respectively. The durational differences between each pair of these sequences were found to be statistically highly significant where p value is at 0.0001 in both tables. These results are shown to be compatible with those of the articulatory timing, viz. T_3, of the same sequences as indicated in tables 6.25 and 6.28. t_3 has been taken to represent $t_1 + t_2$, i.e. the acoustic timing between the end of word-initial [ħ] or [b] and the beginning of word-medial [s] and [ss] or [d] and [dd] plus the acoustic duration of the investigated consonant.

6.7.8 General

From the results presented in all the tables of this experiment, it seems clear that word-initial and word-medial geminate consonants, both plosives and fricatives, are characterized by the following aerodynamic, articulatory and acoustic features:

1. slightly higher oral pressure;
2. smaller volume airflow rate (for fricatives);
3. smaller cross-sectional area (for fricatives);
4. longer articulatory and acoustic durations (i.e. T_2 and t_2); and
5. longer articulatory and acoustic timings (i.e. T_3 and t_3).

On the other hand, the corresponding non-geminate counterparts are characterized by the following features:

1. lower oral pressure;
2. greater volume airflow rate (for fricatives);
3. bigger cross-sectional area (for fricatives);
4. shorter articulatory and acoustic durations (i.e. T_2 and t_2); and
5. shorter articulatory and acoustic timings (i.e. T_3 and t_3).

Apart from the articulatory timing between Ac min. of word-initial [ħ] and the beginning of word-medial [s] and [ss], the tables show that durational differences between vowels preceding geminate consonants and those preceding non-geminate ones were negligible. These differences in duration were statistically shown to be totally non-significant. The tables also illustrate that the standard deviation, i.e. S.D., was generally greater with the geminates than that with the non-geminates. This, undoubtedly, indicates that the dispersion among the measures of geminate consonants was greater than that with their single partners.

Furthermore, a visual scrutiny of the resulting trace patterns does confirm the reliability of the statistical evaluations as well as the results obtained from our segmentation and measurements procedures. However, there is one more thing which, we believe, needs mentioning here. That is, the oral pressure traces are seen to take a relatively small and short domed shape for the single fricative both in word-initial and word-medial positions; while for the geminate cognate the situation seems quite different. The most common shape observable for [ss] in both word positions is either that it takes the pattern of a regular domed curve, or that the traces begin with a slight gradual ascending movement which, when it reaches its farthest end, goes down abruptly indicating the termination of the geminate [ss] and the initiation of the following vowel sound. The other less common noticeable shape is that the oral pressure traces take the reverse pattern when the highest part of the dome comes at the beginning and then the traces take a gradual descending movement where it suddenly ends with a sharp downward movement indicating the end of [ss] and the beginning of the following vowel sound. This is seen when [ss] occurs both word-initially and word-medially.

Table (6.4)

Peak Po (in cm H_2O) for [s] and [ss] used word-initially.

Tokens	[s]	[ss]
1	13.4	10.9
2	14.2	13.9
3	13.1	16.7
4	14.7	16.4
5	15.6	16.1
6	14.5	15.6
7	15.3	16.7
8	14.7	16.4
9	15.3	16.1
10	14.2	15.9
Mean	14.5	15.5
S.D.	0.8	1.8
U	18.5	
P	0.0170*	

Table (6.5)

Peak Po (in cm H_2O) for [s] and [ss] used word-medially.

Tokens	[s]	[ss]
1	13.9	14.2
2	14.7	13.9
3	14.2	15.0
4	15.3	14.5
5	14.2	15.9
6	14.2	15.3
7	15.6	15.9
8	14.5	15.6
9	14.7	15.6
10	13.6	13.9
Mean	14.5	15.0
S.D.	0.6	0.8
U	32.5	
P	0.1822	

Table (6.6)

Peak Po (in cm H_2O) for [d] and [dd] used word-initially.

Tokens	[d]	[dd]
1	11.1	14.5
2	9.2	13.6
3	10.3	12.5
4	11.7	13.9
5	9.5	13.6
6	11.7	13.6
7	11.1	14.2
8	12.0	15.0
9	11.1	13.9
10	11.1	15.6
Mean	10.9	14.0
S.D.	0.9	0.9
U	0	
p	0.0001**	

Table (6.7)

Peak Po (in cm H_2O) for [d] and [dd] used word-medially.

Tokens	[d]	[dd]
1	11.4	13.9
2	11.1	12.0
3	9.7	12.0
4	11.4	13.1
5	10.6	12.2
6	10.6	13.6
7	11.1	13.9
8	11.4	14.2
9	12.0	13.4
10	11.4	13.4
Mean	11.1	13.2
S.D.	0.6	0.8
U	1	
p	0.0002**	

Table (6.8)

Uo (in cm^3/sec) at Ac min. for [s] in /sabit/ and [ss] in /ssabit/.

Tokens	[s]	[ss]
1	424.2	363.6
2	424.2	363.6
3	272.7	272.7
4	393.9	363.6
5	363.6	303.0
6	424.2	303.0
7	484.8	363.6
8	424.2	333.3
9	424.2	303.0
10	545.5	393.9
Mean	418.2	336.3
S.D.	71.2	39.0
U	13	
p	0.0044**	

Table (6.9)

Uo (in cm^3/sec) at Ac min. for [s] in /ħasan/ and [ss] in /ħassan/.

Tokens	[s]	[ss]
1	363.6	303.0
2	333.3	303.0
3	363.6	272.7
4	393.9	212.1
5	393.9	272.7
6	363.6	272.7
7	363.6	333.3
8	454.5	272.7
9	393.9	303.0
10	454.5	363.6
Mean	387.8	290.9
S.D.	39.9	40.9
U	3.5	
p	0.0004**	

Table (6.10)

Po (in cm H_2O) at Ac min. for [s] in /sabit/ and [ss] in /ssabit/.

Tokens	[s]	[ss]
1	13.3	10.6
2	14.2	13.9
3	13.3	16.7
4	14.7	16.1
5	15.6	15.8
6	14.5	15.3
7	15.0	16.1
8	14.7	16.1
9	15.3	16.1
10	14.2	15.8
Mean	14.5	15.3
S.D.	0.8	0.8
U	19.5	
p	0.0204*	

Table (6.11)

Po (in cm H_2O) at Ac min. for [s] in /ħasan/ and [ss] in /ħassan/.

Tokens	[s]	[ss]
1	13.9	14.2
2	14.7	13.9
3	14.2	14.7
4	15.0	14.4
5	14.2	15.6
6	14.2	15.3
7	15.3	15.8
8	14.4	15.6
9	14.7	15.6
10	13.6	13.9
Mean	14.4	14.9
S.D.	0.5	0.8
U	31.5	
p	0.1582	

Table (6.12)

Uo (in cm^3/sec) at Ac min. for [s] in /sabit/ and in /ħasan/.

Tokens	/sabit/	/ħasan/
1	424.2	363.6
2	424.2	333.3
3	272.7	363.6
4	393.9	393.9
5	363.6	393.9
6	424.2	363.6
7	484.8	363.6
8	424.2	454.5
9	424.2	393.9
10	545.5	454.5
Mean	418.2	387.8
S.D.	71.2	39.9
U	30.5	
p	0.1328	

Table (6.13)

Uo (in cm^3/sec) at Ac min. for [ss] in /ssabit/ and in /ħassan/.

Tokens	/ssabit/	/ħassan/
1	363.6	303.0
2	363.6	303.0
3	272.7	272.7
4	363.6	212.1
5	303.0	272.7
6	303.0	272.7
7	363.6	333.3
8	333.3	272.7
9	303.0	303.0
10	393.9	363.7
Mean	336.3	290.9
S.D.	39.0	40.9
U	21	
p	0.0240*	

Table (6.14)

Po (in cm H_2O) at Ac min. for [s] in /sabit/ and in /ħasan/.

Tokens	/sabit/	/ħasan/
1	13.3	13.9
2	14.2	14.7
3	13.3	14.2
4	14.7	15.0
5	15.6	14.2
6	14.5	14.2
7	15.0	15.3
8	14.7	14.4
9	15.3	14.7
10	14.2	13.6
Mean	14.5	14.4
S.D.	0.8	0.5
U	45	
p	0.7019	

Table (6.15)

Po (in cm H_2O) at Ac min. for [ss] in /ssabit/ and in /ħassan/.

Tokens	/ssabit/	/ħassan/
1	10.6	14.2
2	13.9	13.9
3	16.7	14.7
4	16.1	14.4
5	15.8	15.6
6	15.3	15.3
7	16.1	15.8
8	16.1	15.6
9	16.1	15.6
10	15.8	13.9
Mean	15.3	14.9
S.D.	1.8	0.8
U	24.5	
p	0.0518	

Table (6.16)

Ac min. $(kU / \sqrt{\Delta P})$ in cm^2 for [s] in /sabit/ and [ss] in /ssabit/.

Tokens	[s]	[ss]
1	0.088	0.085
2	0.086	0.074
3	0.057	0.051
4	0.078	0.069
5	0.070	0.058
6	0.085	0.059
7	0.095	0.069
8	0.084	0.063
9	0.082	0.057
10	0.110	0.075
Mean	0.084	0.066
S.D.	0.014	0.010
U	15	
p	0.0081**	

Table (6.17)

Ac min. $(kU / \sqrt{\Delta P})$ in cm^2 for [s] in /ħasan/ and [ss] in /ħassan/.

Tokens	[s]	[ss]
1	0.074	0.061
2	0.066	0.062
3	0.073	0.054
4	0.077	0.042
5	0.079	0.052
6	0.073	0.053
7	0.071	0.064
8	0.091	0.052
9	0.078	0.058
10	0.094	0.074
Mean	0.078	0.057
S.D.	0.009	0.009
U	4.5	
p	0.0006**	

Table (6.18)

Ac min. $(kU / \sqrt{\Delta P})$ in cm^2 for [s] in /sabit/ and in /ħasan/.

Tokens	/sabit/	/ħasan/
1	0.088	0.074
2	0.086	0.066
3	0.057	0.073
4	0.078	0.077
5	0.070	0.079
6	0.085	0.073
7	0.095	0.071
8	0.084	0.091
9	0.082	0.078
10	0.110	0.094
Mean	0.084	0.078
S.D.	0.015	0.009
U	36	
p	0.2694	

Table (6.19)

Ac min. $(kU / \sqrt{\Delta P})$ in cm^2 for [ss] in /ssabit/ and in /ħassan/.

Tokens	/ssabit/	/ħassan/
1	0.085	0.061
2	0.074	0.062
3	0.051	0.054
4	0.069	0.042
5	0.058	0.052
6	0.059	0.053
7	0.069	0.064
8	0.063	0.052
9	0.057	0.058
10	0.075	0.074
Mean	0.066	0.057
S.D.	0.010	0.009
U	38	
p	0.3367	

Table (6.20)

Articulatory duration (T_2) in msec, of the supraglottal constriction for [s] and [ss] used word-initially.

Tokens	[s]	[ss]
1	100	215
2	80	200
3	80	170
4	80	180
5	70	170
6	80	170
7	80	165
8	80	180
9	80	170
10	80	195
Mean	80	180
S.D.	7.9	16.5
U	0	
p	0.0001**	

Table (6.21)

Articulatory duration (T_2) in msec, of the supraglottal constriction for [s] and [ss] used word-medially.

Tokens	[s]	[ss]
1	60	180
2	60	185
3	45	170
4	55	140
5	55	130
6	55	195
7	55	130
8	55	130
9	45	165
10	55	130
Mean	55	155
S.D.	5.2	26.2
U	0	
p	0.0001**	

N.B. The numbers in the above tables are rounded up to the nearest 5 msec because the differences between the two sets are very clear-cut.

Table (6.22)

Articulatory duration (T_2), in msec, of the supraglottal constriction for [d] and [dd] used word-initially.

Tokens	[d]	[dd]
1	70	210
2	40	195
3	50	170
4	50	170
5	50	155
6	50	165
7	30	165
8	50	155
9	50	130
10	50	200
Mean	50	170
S.D.	10.0	24.1
U	0	
p	0.0001**	

Table (6.23)

Articulatory duration (T_2), in msec, of the supraglottal constriction for [d] and [dd] used word-medially.

Tokens	[d]	[dd]
1	60	185
2	55	195
3	45	170
4	55	210
5	40	170
6	45	170
7	45	180
8	45	185
9	40	155
10	45	165
Mean	50	180
S.D.	6.8	16.0
U	0	
p	0.0001**	

N.B. The numbers in the above tables are rounded up to the nearest 5 msec because the differences between the two sets are very clear-cut.

Table (6.24)

Articulatory timing (T_1), in msec, between Ac min. of [ħ] and the initiation of [s] in /ħasan/ and of [ss] in /ħassan/.

Tokens	/ħasan/	/ħassan/
1	108.5	124.0
2	100.8	124.0
3	108.5	108.5
4	108.5	108.5
5	100.8	124.0
6	93.0	116.3
7	100.8	116.3
8	93.0	124.0
9	108.5	131.8
10	93.0	116.3
Mean	101.5	119.4
S.D.	6.8	7.5
U	4	
p	0.0004^{**}	

Table (6.25)

Articulatory timing (T_3), in msec, between Ac min. of [ħ] and the termination of [s] in /ħasan/ and of [ss] in /ħassan/.

Tokens	/ħasan/	/ħassan/
1	168.5	304.0
2	160.8	309.0
3	153.5	278.5
4	163.5	248.5
5	155.8	254.0
6	148.0	311.3
7	155.8	246.3
8	148.0	254.0
9	153.5	296.8
10	148.0	246.3
Mean	155.5	274.9
S.D.	7.0	28.0
U	0	
p	0.0001^{**}	

Table (6.26)

Aritculatory timing (T_4), in msec, between Ac min. of [ħ] and Ac min. of [s] and [ss] in /ħasan/ and /ħassan/, respectively.

Tokens	/ħasan/	/ħassan/
1	139.5	232.6
2	139.5	255.8
3	139.5	224.8
4	131.8	209.3
5	131.8	217.1
6	124.0	217.1
7	131.8	217.1
8	124.0	209.3
9	131.8	224.8
10	124.0	217.1
Mean	131.8	222.5
S.D.	6.3	13.7
U	0	
p	0.0001**	

Table (6.27)

Articulatory timing (T_1), in msec, between occlusion offset of [b] and occlusion onset of [d] in /badal/ and [dd] in /baddal/.

Tokens	/badal/	/baddal/
1	85.3	77.5
2	77.5	77.5
3	85.3	85.3
4	85.3	77.5
5	77.5	77.5
6	77.5	77.5
7	85.3	77.5
8	85.3	85.3
9	77.5	85.3
10	77.5	85.3
Mean	81.4	80.6
S.D.	4.1	4.0
U	45	
p	0.6613	

Table (6.28)

Articulatory timing (T_3), in msec, between occlusion offset of [b] and occlusion offset of [d] and [dd] in /badal/ and /baddal/, respectively.

Tokens	/badal/	/baddal/
1	145.3	262.5
2	132.5	272.5
3	130.3	255.3
4	140.3	287.5
5	117.5	247.5
6	122.5	247.5
7	130.3	257.5
8	130.3	270.3
9	117.5	240.3
10	122.5	250.3
Mean	128.9	259.1
S.D.	9.2	14.3
U	0	
p	0.0001**	

Table (6.29)

Acoustic duration (t_2), in msec, for [s] and [ss] used word-initially.

Tokens	[s]	[ss]
1	131.8	248.1
2	124.0	217.1
3	112.4	209.3
4	100.8	193.8
5	108.5	186.0
6	115.0	201.6
7	116.3	178.3
8	108.5	197.7
9	116.3	193.8
10	116.3	209.3
Mean	115.0	203.5
S.D.	8.6	19.5
U	0	
p	0.0002^{**}	

Table (6.30)

Acoustic duration (t_2), in msec, for [s] and [ss] used word-medially.

Tokens	[s]	[ss]
1	108.5	232.6
2	104.7	217.1
3	100.8	170.5
4	100.8	182.2
5	100.8	174.4
6	96.9	220.9
7	96.9	174.4
8	85.3	178.3
9	93.0	201.6
10	85.3	158.9
Mean	97.3	191.1
S.D.	7.6	25.1
U	0	
p	0.0002^{**}	

Table (6.31)

Acoustic duration (t_2), in msec, for [d] and [dd] used word-initially.

Tokens	[d]	[dd]
1	62.0	201.6
2	31.0	182.2
3	31.0	162.8
4	34.9	155.0
5	31.0	139.5
6	38.8	147.3
7	31.0	155.0
8	46.5	151.2
9	38.8	124.0
10	46.5	186.0
Mean	39.2	160.5
S.D.	10.1	23.4
U	0	
p	0.0001**	

Table (6.32)

Acoustic duration (t_2), in msec, for [d] and [dd] used word-medially.

Tokens	[d]	[dd]
1	46.5	178.3
2	46.5	155.0
3	38.8	162.8
4	46.5	201.6
5	34.9	162.8
6	46.5	162.8
7	38.8	166.7
8	34.9	182.2
9	34.9	151.2
10	34.9	162.8
Mean	40.3	168.6
S.D.	5.5	14.9
U	0	
p	0.0001**	

Table (6.33)

Acoustic timing (t_1), in msec, between the termination of [ħ] and the initiation of [s] and [ss] in /ħasan/ and /ħassan/, respectively.

Tokens	/ħasan/	/ħassan/
1	62.0	62.0
2	54.3	69.8
3	54.3	62.0
4	58.1	58.1
5	54.3	62.0
6	62.0	54.3
7	62.0	62.0
8	62.0	69.8
9	54.3	54.3
10	54.3	50.4
Mean	57.8	60.5
S.D.	3.8	6.4
U	37.5	
p	0.3174	

Table (6.34)

Acoustic timing (t_1), in msec, between the termination of [b] and the initiation of [d] and [dd] in /badal/ and /baddal/, respectively.

Tokens	/badal/	/baddal/
1	85.3	85.3
2	77.5	85.3
3	85.3	81.4
4	81.4	81.4
5	81.4	81.4
6	77.5	77.5
7	85.3	89.1
8	89.1	93.0
9	85.3	93.0
10	81.4	85.3
Mean	83.0	85.3
S.D.	3.7	5.2
U	38	
p	0.3461	

Table (6.35)

Acoustic timing (t_3), in msec, between the termination of [ħ] and the termination of [s] and [ss] in /ħasan/ and /ħassan/, respectively.

Tokens	/ħasan/	/ħassan/
1	170.5	294.6
2	159.0	286.9
3	155.1	232.5
4	158.9	240.3
5	155.1	236.4
6	158.9	275.2
7	158.9	236.4
8	147.3	248.1
9	147.3	255.9
10	139.6	209.3
Mean	155.1	251.6
S.D.	8.5	26.1
U	0	
p	0.0001**	

Table (6.36)

Acoustic timing (t_3), in msec, between the termination of [b] and the termination of [d] and [dd] in /badal/ and /baddal/, respectively.

Tokens	/badal/	/baddal/
1	131.8	263.6
2	124.0	240.3
3	124.1	244.2
4	127.9	283.0
5	116.3	244.2
6	124.0	240.3
7	124.1	255.8
8	124.0	275.2
9	120.2	244.2
10	116.3	248.1
Mean	123.3	253.9
S.D.	4.8	15.2
U	0	
p	0.0001**	

6.8 Main Experiment

6.8.1 Introduction and Aims

It is only over the last two decades or so that researchers have began to map out what happens, acoustically and articulatorily, when a speaker alters his speaking rate. Research is being increasingly directed towards finding acoustic and articulatory invariance across changes in rate of speech. Some researchers (e.g. Lehiste, 1972, 1973; Oller, 1973; Klatt, 1975; Cooper, 1976) have reported that the measured duration of a given linguistic segment varies as a function of its placement in a larger linguistic unit, no matter whether this unit is a sylalble, a word or a phrase. For instance, a linguistic segment occurring within such a linguistic unit is very often longer in duration when it terminates the unit than when it occurs in the middle of it. Initial segments in a larger linguistic unit have been found to be longer than medial ones; final segments are the longest of all.

Other factors that may affect the duration of a segment have also been reported. Fowler (1977), for example, reports that these factors are the flexibility in rate of speaking that a talker may exploit to fit different conversational circumstances, and the different levels of stress that are evident in all utterances. On the other hand, Gaitenby (1965) and Kent et al. (1974) found that both the segment durations measured from the acoustic record of an utterance and the extent of articulatory movements decreased proportionately with an increase in rate of speech. Fowler (op.cit.) points out that Gaitenby (op.cit.) compared the utterances of fast and slow talkers and reported the greatest durational differences between fast and slow versions of stressed vowels. She also states that "voiced consonants reduced the least with an increase in speech rate." [p.23].

At the same time, studies on vowel sounds show that a reduction in the duration of a vowel is accompanied by a concomitant reduction in its spatial coordinates. Lindblom (1963), for instance, reports that the articulatory movements towards the required shapes of the vocal tract for the vowel are increasingly less extensive as speaking rate increases. He suggests that this vowel reduction is caused by a change in the rate at which commands to the articulators are released. He adds that the target value of a vowel is invariant over different speaking rates, but that it is not reached at a fast speaking rate because the command for the next sound is actualized before the target value for the vowel is attained. Because vowel reduction appeared for changes in both stress and speech rate in Lindblom's data, he concluded that reduction and undershoot were caused solely by changes in duration, and not the suprasegmental features of stress and speaking rate, per se. Lindblom's original hypothesis of undershoot was proposed to account for changes in both stress and speaking rate. That is, his model predicts that undershoot would occur for both destressed and faster speech. (Gay and Hirose, 1973).

Though researchers (e.g. Kozhevnikov and Chistovich, 1965; Klatt, 1976) contend that vowels shorten more than consonants when speaking rate increases, yet Gaitenby's data (op.cit.) on the production of slow and fast speakers illustrate that consonants shorten as well as vowels at fast rates of speech. In their investigation, Kozhevnikov and Chistovich observed that there is a distinct relation between the relative duration of the consonant and vowel in a syllable and the speaking rate. They found that a vowel can disappear completely with a rapid rate of speech. Their findings indicate that variations of speaking rate significantly change the relative durations of the consonants and vowels within syllables and that, in the case of a rather slow speaking rate, the duration of the consonant "practically does not change and the prolongation of the syllable occurs at the expense of the vowel." [p.89].

Fitch (1981) designed three experiments to investigate temporal differences between intervocalic voiced and voiceless plosive consonants, and to pursue the distinction between temporal information for voicing and temporal information for rate. He claims that these differences can be produced by changes in speaking rate as well as by closure duration of the segment and the vowel duration preceding the closure. His experiments aimed to confirm that both closure duration and vowel duration can cue the intervocalic phonemic voicing difference between words, such as /dabi/ and /dapi/, and that both closure duration and vowel duration can cue the rate difference between fast and slow speech.

He states that "a voicing change from [b] to [p] lengthens the closure and shortens the vowel; a rate change from fast to slow lengthens both." [p.1]. It was found that the relationship between closure duration and vowel duration greatly affected judgments of rate but not those of voicing. Fitch also observed that the relationship between the duration of the initial consonant transitions and the duration of the steady-state vowel nucleus was perceptually salient, and might serve to distinguish information for voicing from information for rate.

On the other hand, electromyographic data on stop-vowel sequences for five muscles associated with lip, tongue and jaw movements presented by Tuller et al. (1981a, 1981b) indicated that some consonant-related activity and some vowel-related activity showed fixed relative timing, preserved across a change in rate and syllable stress. Tuller et al. (1981a) explored the temporal relations among electromyographic measures of articulatory events, and the pattern of changes in individual muscle actions as a function of syllable stress and speaking rate. Their results indicated that "large variations were found in the magnitude and duration of activity in each muscle; variations accompanying speaking rate change were not equivalent to the variations accompanying a change in stress." [p.33]. These results did not support the notion that acoustic shortening, which accompanied both decreases in syllable stress and increases in speaking rate in their experiments, was the product of a single style of articulatory change. Tuller et al. (1981b) continued their earlier work and designed an experiment to explore the possibility that relative timing of articulatory events is preserved over suprasegmental change. Electromyographic recordings from lip, tongue and jaw muscles were obtained during production of utterances whose phonetic structure allowed intersegmental timing relationships to be examined over more than two phonetic segments. Their findings show that the preservation of relative timing of muscle activity over metrical change is characteristic of the temporal organization of speech. On these findings, Scully (1983) comments, and states that "this pattern of muscle forces could perhaps result in local invariance of inter-articulator co-ordination." [p.11]. Fujimura (1983) has recently found that, in the case of stop-vowel sequences, the relatively invariant 'iceberg pattern' of articulation is expected to control crucial formant transitions for consonant and vowel identification.

Further, investigations into the production of syllable-initial consonants at different rates of articulation have revealed changes in the speed with which the relevant gestures are made and also in the resulting acoustic transitions (Gay, 1978; Gay and Hirose, 1973; Gay et al., 1974; Soli, 1982). If one takes into account that changes in the duration of a transition occur with changes in rate, one should expect that the listener would make the appropriate normalization when using transition duration to cue a phonetic distinction. (Miller and Liberman, 1979). For instance, Gay (ibid.) has reported that the initial transitions for [b], in intervocalic position, are somewhat longer at slower speech rates. Soli (op.cit.), similarly, has shown that the initial transitions for the semi-vowel [j], in syllable-initial position, lengthen considerably as speech is slowed and the syllable becomes longer.

Besides, results of electromyographic studies on the effects of speaking rate on labial consonants provide some interesting support for the acoustic stability of the stop-glide contrast. Mack and Blumstein (1983) remark that these studies have shown that speaking rate affects the production of stops and glides in a similar manner. That is, the targets for lip opening and closing are essentially the same across different rates of speech. They also state that "the characterizing differences in electromyographic patterns of labial stops and glides are maintained even as their production rate changes. So the stability of articulatory measures suggests that there may be a consequent acoustic invariance differentiating stops from glides." [p.1749].

In a series of experiments carried out by Gay and his co-workers (e.g. Gay, 1968; Gay, 1978; Gay, 1981a, 1981b; Gay and Hirose, 1973; Gay et al., 1974; Gay and Ushijima, 1975) it was shown that the motor patterns underlying articulatory movements for fast speech were not only different than those during slow speech, but were also reorganized in complex ways. For example, it was documented that electromyographic activity associated with tongue body movements during vowel production generally decreased during fast speech;

whereas activity associated with both labial and alveolar plosive consonant production increased with an increase in speaking rate. Gay and Hirose (op.cit.) reported that the major finding of their investigation indicated that an increase in speaking rate was accompanied by both an increase in the activity level of the muscle as well as an increase in the rate of movement of the lips. These differences in muscle activity levels denote that the control of speaking rate requires more than just a simple adjustment of the timing of motor commands, and that although changes in the timing of commands to the muscles do occur, the production of labial consonants during faster speech is characterized primarily by an increase in articulatory effort. Furthermore, Gay and Ushijima (op.cit.) stated that for lip movement associated with either labial consonant production or vowel rounding, and for tongue tip movement associated with lingual consonant production, an increase in speaking rate was accompanied by an increase in the activity level of the muscle. On the other hand, for tongue movement during vowel production, an increase in speaking rate had the opposite effect. These electromyographic data were found to be inconsistent with the acoustic data documented by Gay (1978) who reported that differences in vowel duration due to changes in speaking rate did not seem to have a substantial effect on the attainment of acoustic vowel targets. He also added that the formant frequencies of these presumed targets remained essentially unchanged across changes in speaking rate.

Kelly et al. (1962) presented acoustic evidence that the slopes of F2 transitions might be characteristic of particular demisyllables used in sentences spoken at two different tempi by four English native speakers. Their work concentrated on F2 in sequences of labial consonant plus vowel partly because they believed that this combination frequently shows a good deal of movement, and partly because they were interested in the transitional differences visually associated with the [b] ~ [w] distinction. Despite the dearth of evidence for F2 transitions for [b] and the limitation of their test items, their findings indicated that, for a given vowel following [w], the rate of F2 transitions remained unchanging at different speaking rates. The evidence was not clear on the rate of [b] transitions at different tempi. It was also found that [b] and [w] followed by the same vowel had identical F2 transition rates and differed only in onset position. In a subsequent investigation, Miller and Liberman (op.cit.) showed that the duration of the initial formant transition sufficient to differentiate perceptually the plosive consonant [b] from the semivowel [w] was not absolute. They found that as the overall syllable duration was lengthened, an increasingly relatively longer transition was required to perceive /wa/, as opposed to /ba/. They interpreted this effect as an adjustment for variations in speech rate and contended that the effect of transition duration as a cue for the [b] ~ [w] distinction is influenced by the duration and structure of the syllable containing the cue as well as by the duration of a following syllable. In their opinion "this after-going effect reflects an adjustment by the listener to the articulatory rate of the speaker: the duration and structure of the syllable provide information about rate, and the listener uses this information when making a phonetic judgment of [b] vs. [w]." [p.464]. Therefore, as overall syllable duration increased, due to a decrease in speaking rate, the transition duration that distinguished [b] and [w] increased.

In a more recent investigation, Miller and Baer (1983) have examined the way in which the transition durations of [b] and [w] change as a function of speaking rate. They found that, at any given rate of speech, the /wa/ transitions were longer than the /ba/ transitions, and that the magnitude of the difference between average /ba/ and /wa/ transition durations increased with decrease in speech rate. There was a considerable increase in the initial transition duration of /wa/; while there was little change in the initial transition duration of /ba/. Miller and Baer remarked that their findings were in agreement with those reported by Miller and Liberman (op.cit.) in that when listeners identify /ba/ and /wa/ on the basis of transition duration, they do so in relation to the duration of the syllable. They concluded that "there is a close correspondence between the way in which alterations in speaking rate affect the durations of transition for [b] and [w] and the way in which listeners use this acoustic property to distinguish the two segments across a range of speaking rates." [p.1754]. These findings, Mack and Blumstein (op.cit.) comment, suggest that absolute duration cannot provide an invariant cue to the stop-glide contrast and, more importantly, that this contrast seems to be context-dependent, i.e. dependent on syllable duration and thus presumably on the subject's speaking rate.

Following earlier investigations (e.g. Blumstein and Stevens, 1979, 1981; Stevens and Blumstein, 1978, 1981), Mack and Blumstein (op.cit.) conducted a comparative acoustic analysis to identify an invariant acoustic property which could distinguish the plosive [b] and the semi-vowel [w] in five vowel contexts by two speakers. Their results showed that transition durations and formant frequencies often differ considerably in these two types of consonants. According to them, such differences might suggest that either a frequency or duration measure contributes to an invariant property for distinguishing stops and glides. This fact brought them to the conclusion that it was impossible to extract an invariant property for distinguishing between these two types of sounds merely on the basis of duration or frequency measures. It was also found that the relative amplitude change was systematically larger in the vicinity of the [b] release than of the [w] release across vowel contexts and speakers. These findings made Scully (op.cit.) comment that "If the reliability of the contrast is preserved across rate changes, then this could be, as they suggest, an invariant property characterising stop-glide (or approximant) contrasts." [p.11].

After this brief review of the most relevant works that dealt with speech rate and its effect on the production and perception of speech sounds, it is worth stating that the aim of the present experiment[1] was to go a step beyond our pilot experiment to examine the aerodynamic, articulatory and acoustic aspects of speech production while saying single and geminate consonants in I.C. Arabic. It was hoped that this experiment would:

1. discover the aerodynamic, articulatory and acoustic invariances across changes in rate of speech;
2. discover the differences and similarities in speech among native speakers of one and the same dialect; and
3. examine the differences and similarities that might arise between utterances spoken with slow (i.e. normal) speaking rate and the same utterances spoken with a faster speaking rate.

6.9 Experimental Procedure

6.9.1 Selection of Stimuli and Recording Technique

The stimulus items selected for the present main experiment were the same as those referred to in our previous experiments, see section 3.2.1. The chosen words were used in the same constructed carrier sentence, viz. /'kitbi____'sit mar'raat/, thus retaining the same accentual pattern.

A list comprising 48 items, 6 tokens for each word, was first recorded by the experimenter in the recording studio of the Phonetics Laboratory. The 48 items were spoken

---

[1] Also see section 6.12.1.

with two different tempi. They were first uttered with a slow[1] tempo and then with a faster tempo, and therefore making a total of 96 items. There were six runs in all; three runs spoken with a slow speaking rate and the other three with a faster rate. Each run contained sixteen items. At the beginning of each run the dummy carrier sentence /ˈkitbi ki'taab ˈsit marˈraat/ was pronounced. The dummy sentence was excluded when segmentation and measurements were made later.

Using a computer program, called OKPROMPT, the recorded items were randomized anew in each of the six runs with the help of a micro-computer (type Commodore, PET 2001 series) which was placed in front of the experimenter inside the recording studio while the recordings were performed. It was the computer program that helped the experimenter to select the appropriate period of time for the recorded items to be displayed as well as the blank time value needed between the appearance of one item and the next on the computer screen. After several initial attempts, it was then found that in the case of the slow speaking rate it was convenient to have 1.5 sec as 'the time value for screen display' and 2 sec as 'the blank time value between items'; whereas in the case of the faster speech rate it was found that 1 sec and 1.75 sec were convenient as 'the time value for screen display' and as 'the blank time value between items', respectively.

---

1. By 'slow' we mean 'normal'. 'Slow' is used here as opposed to 'fast'.

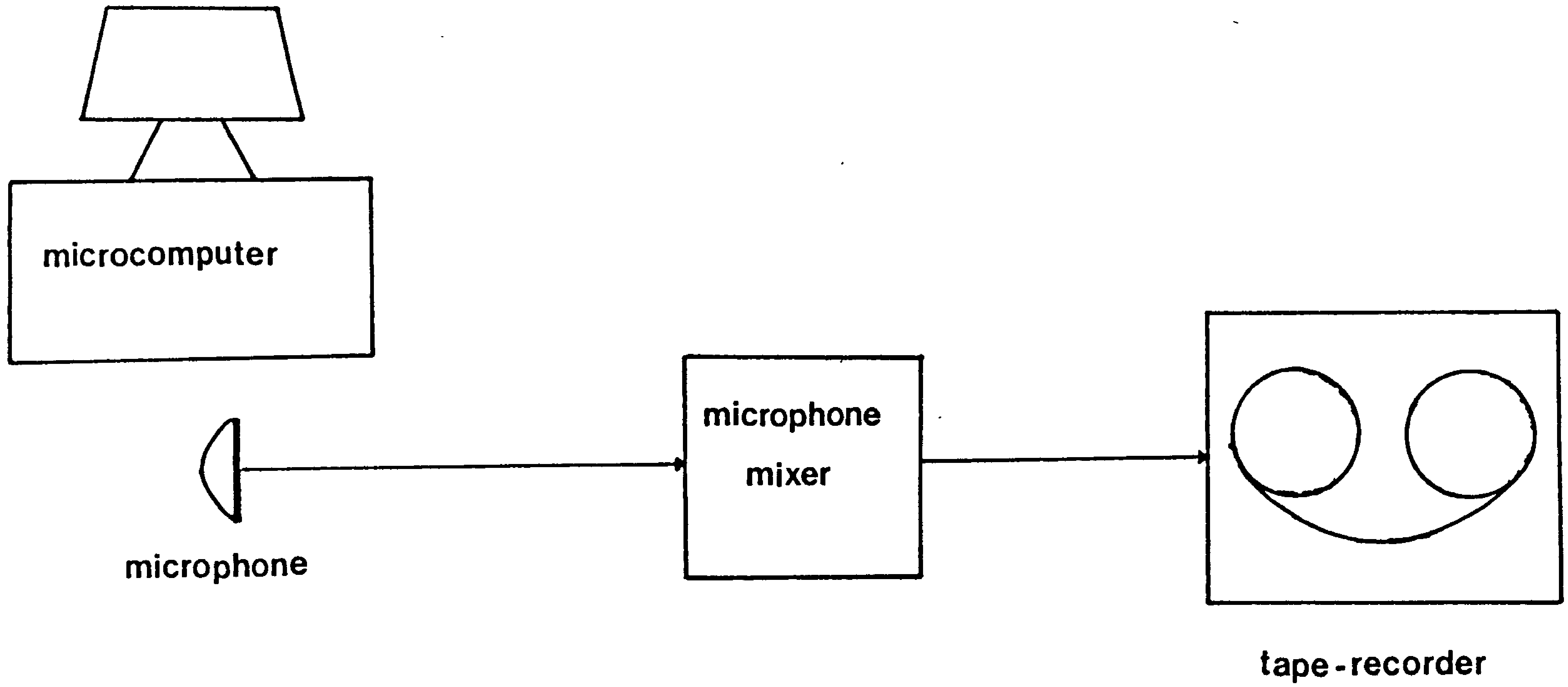
The stimulus items were recorded on high quality Ampex tape by means of a stereo tape-recorder (type Revox A77). The tape recording speed was 7½"/sec, and the microphone used was the AKG (type D202 E1). The distance from the speaker to the microphone was 13". In order to amplify the signal that goes from the microphone to the tape-recorder, a microphone mixer (type Shure Professional) was employed. The gain of the microphone mixer was set at 3.

It was not necessary to repeat the same recording procedure while making the studio recording for the other subject. The recorded stimuli, spoken by the experimenter, were played back on the Revox A77 tape-recorder and were fed into the ears of the second subject via very high quality Electrostatic Headphones (type Koss/ESP6A). By employing this technique, it was hoped that the second subject would retain the same speaking rates and the same accentual patterns already used by the experimenter. Having identical recording conditions would perhaps minimize the interference of extraneous variables and, therefore, make the final results of the two subjects comparable.

The items spoken by the second subject were also recorded on an Ampex tape using the Revox A77 tape-recorder. The speed of the tape was 7½"/sec. The same AKG microphone (type D202E1) was employed. The distance from the speaker to the microphone was 10", and the gain of the microphone mixer was set at 4.

The tape recordings were later used as playbacks while conducting the main experiment by the two subjects in

Fig.(6.10) Block diagram of instrumental set-up used for studio recordings



the Phonetics Laboratory. Fig. 6.10 illustrates the instrumentation used for the studio recordings in this experiment.

6.9.2 Subjects

The author of the present study (subject GG) and another male post-graduate student (subject AM) acted as two subjects in this experiment. They are both native speakers of I.C. Arabic with a clear prevailing accent of Basrah dialect.[1]

Subject AM was about 31 years old. He had spent most of his previous life in the city of Basrah and he had been educated there. He got his first university degree in English from the University of Basrah where he was later appointed as a research assistant. In 1979, he came to Leeds where he joined the Department of Linguistics and Phonetics to get a higher degree in Linguistics.

In regard to experimental phonetics, subject AM was a complete novice. He had had no opportunity to participate in an experiment, such as the present one, before.

6.9.3 Instrumental Set-Up

6.9.3.1 Airflow

The same rubber facemask exploited in our pilot experiment was used by subject GG in the present investigation.

---

[1] For further information about the linguistic background of subject GG, see section 3.2.2.

The mask also fitted well on the face of subject AM. For each subject a preliminary checking had been made before starting the main experiment by breathing in and out first through the nose and then through the mouth to make sure that there was no escape of air through the rigid rubber flange that separates the nose division from the mouth division inside the mask. For both subjects the nasal airflow traces and the oral airflow traces were graphically recorded on Channels 1 and 2 of the mingograph, respectively. Calibration of the oral airflow traces was accomplished by using a rotameter. The airflow values are in cm^3/sec.

6.9.3.2 Oral Pressure

For the purpose of obtaining oral pressure traces, two different methods were used by the subjects. While it was convenient for subject GG to continue using the soft polyethylene tube (type Franklin's oesophageal tube, size 8 ch-257) inserted through his nose down to the pharynx, subject AM found it rather inconvenient and irritating to insert a similar tube, or even one of a smaller size, through his nose. He made several attempts to carry out the experiment with the tube through his nose but he was unsuccessful. Accordingly, the soft tube was replaced by a more solid and slightly larger one (type WSP 6228, oesophageal stomach tube, 10FG K64). The tube was first heated in boiling water and then bent to fit along the maxillary dental arch, around the last molar and into the posterior midline of the oral cavity. It is assumed that in this position it is least likely to give spurious pressure readings (Hardy, 1965; Scully, 1971). The compressed air behind the constriction was transmitted via the sensing polyethylene tube to a Gaeltec pressure transducer (type 3CT, ± 10 cm H_2O) through the open end inside the mouth.

At the beginning, subject AM experienced some difficulty in placing the pressure tube in the right position when he first inserted it into his mouth. After some unsuccessful attempts, he then managed to put the tube in its proper position where satisfactory results were obtained. The experiment was then conducted. The tube had to be cleared every two runs lest the accumulated saliva and mucus should block it. The oral pressure traces of each subject were displayed on channel 3 of the mingograph. Calibration of the oral pressure traces was made in cm H_2O against a U-tube water manometer.

6.9.3.3 Area

For measuring the timing of the minimum constriction area formed by the tip or blade of the tongue and the alveolar region of the palate in the production of the two fricatives [s] and [ss], the same aerodynamic method reported in our pilot experiment was employed, see section 6.3.3.3.

6.9.3.4 Intensity Level

The acoustic data for the same stimuli were simultaneously obtained by using an intensity meter (F-J Electronics). Two types of intensity level traces were recorded; low pass filtered at 500 Hz (linear scale) and high pass filtered at 3.9 kHz (log scale). Mingograph traces of the high pass filter were obtained by linking the intensity meter to a high pass 3.9 kHz audio-frequency filter (type JSRUR 1) which received and sent signals from and to the intensity meter which, in turn, sent the signals to the mingograph. For the low pass filter the intensity meter was set at 12 dB for subject GG and at 24 dB for subject AM. For the high pass filter it was set at 28 dB for both subjects.

A miniature microphone (type Knowles Electret) was mounted on the facemask inside the mouth space. The signal from this microphone was fed into the Revox A77 tape-recorder used with an Ampex tape (type 1200). The output signal from the tape-recorder was fed into the intensity meter, with the tape-recorder played at a speed of 7½"/sec. The recording was made on two channels. The input to channel 1 was set at 1¾ for subject GG and at 1- for subject AM; whereas the input to channel 2 was set at 5 for the two subjects. The output was stereo set at 7¾ for subject GG and at 5¼ for subject AM.

The low pass intensity level traces filtered at 500 Hz and the high pass intensity level traces filtered at 3.9 kHz were graphically recorded on channels 5 and 6 of the mingograph, respectively.

6.9.3.5 Oscillogram

Ordinary oscillograms of the whole set of material used in this experiment were obtained by attaching a small

Fig.(6.11) **Block diagram of instrumental set-up used in the main experiment to investigate the aerodynamics of single and geminate consonants**

Fig. (6.11)

Description

A	Flowhead for nose airflow (Mercury).
A'	Flowhead for mouth airflow (Mercury).
B	Pressure transducer for nose flowhead (Gaeltec).
B'	Pressure transducer for mouth flowhead (Gaeltec).
C	Pressure transducer for pressure tubes (Gaeltec).
D	Stereo tape-recorder (Revox A77).
E	Control unit for airflow (Gaeltec).
F	Control unit for oral pressure (Gaeltec).
G	Intensity meter (F-J Electronics).
H	High pass 3.9 kHz audio-frequency filter (JSRUR 1).
I	Aerodynamic speech analyzer (Richard Caley).
J	Attenuator (A64, Advance)
K	Mingograph (Mingograf No. 803/Siemens Elema).

Mingograph Traces:

Ch.0	Timer.
Ch.1	Volume flow rate of air through the nose, gain set at 7¼ for both subjects.
Ch.2	Volume flow rate of air through the mouth, gain set at 5 for GG and at 6 for AM.
Ch.3	Oral pressure with air in front of the mouth as reference, gain set at 4½ for GG and at 6 for AM.
Ch.4	'Area' of oral constriction derived from oral airflow and oral pressure, gain set at 6 for GG and at 3 for AM.
Ch.5	Intensity level trace low pass filtered at 500 Hz (linear scale), gain set at 2 for GG and at 5¾ for AM.
Ch.6	Intensity level trace high pass filtered at 3.9 kHz (log scale), gain set at 7 for GG and at 7¼ for AM.
Ch.7	Ordinary oscillogram, gain set at 7½ for GG and at 6 for AM.

microphone to the subject's neck, close to his larynx. The signal from the microphone was fed into the Revox A77 tape-recorder. In order to reduce the strength of the signal that went to the mingograph, the tape-recorder was linked to an attenuator (type Step A64, Advance) which, in turn, sent the signal to the mingograph. Oscillograms were displayed on channel 7 of the mingograph for both subjects.

Fig. 6.11 shows the complete instrumentation employed in the main experiment to investigate the aerodynamics of geminate consonants in I.C. Arabic.

6.10 Segmentation and Measurements Criteria

By and large, the criteria employed in this experiment for segmenting and measuring the investigated stimuli were the same as those used in our pilot experiment on the aerodynamic and articulatory activities of the vocal tract during speech production. Articulatory durations (in msec) were similarly based on establishing points of time on the 'area' trace representing two events corresponding to the initiation and termination of the articulatory occlusion of the two fricatives [s] and [ss], and of the two plosives [d] and [dd] occurring word-initially and word-medially, see sections 6.4.2 and 6.4.3 for full details. Oral pressure values were measured (in cm H_2O) between the baseline and the peak points of the oral pressure trace, see section 6.4.4 for further information.

One of the main problems encountered in the present investigation was in measuring articulatory durations and the relative magnitudes of the minimum values of constriction area for the geminate [ss]. The 'area' trace as well as the oral airflow trace were, sometimes, seen to have two dips instead of one dip as usually seen on the two traces in the case of the non-geminate [s]. This problem was not clearly noticeable in our pilot experiment although on a very few occasions similar 'area' traces were observed, yet they did not seem to invalidate our segmentation and measurements procedures. Such double dips on the oral airflow and 'area' traces were clearly visible in those traces obtained by subject AM. The two traces (i.e. the oral airflow and the 'area' traces) were seen to respond in harmony. That is, they took identical trace patterns.

To solve this problem, the lower dip was chosen for our present articulatory segmentation and measurements criteria whenever the two dips appear on the 'area' trace. The plausibility of our choice of that dip was examined by following a procedure where two points, a and b, were chosen on the 'area' trace, one on each dip, for the geminate [ss] pronounced in word-initial and word-medial positions, as shown in Fig. 6.12. Oral airflow and oral pressure were measured at these two points. Relative magnitudes of constriction area were calculated at the two points, and it was found that the lower point, namely point a, yielded the smallest value of cross-sectional area of constriction in both word positions, see table 6.37.

At the same time, it was found that this choice of the lower dip did not invalidate our procedure in measuring articulatory duration for fricatives. The arbitrary

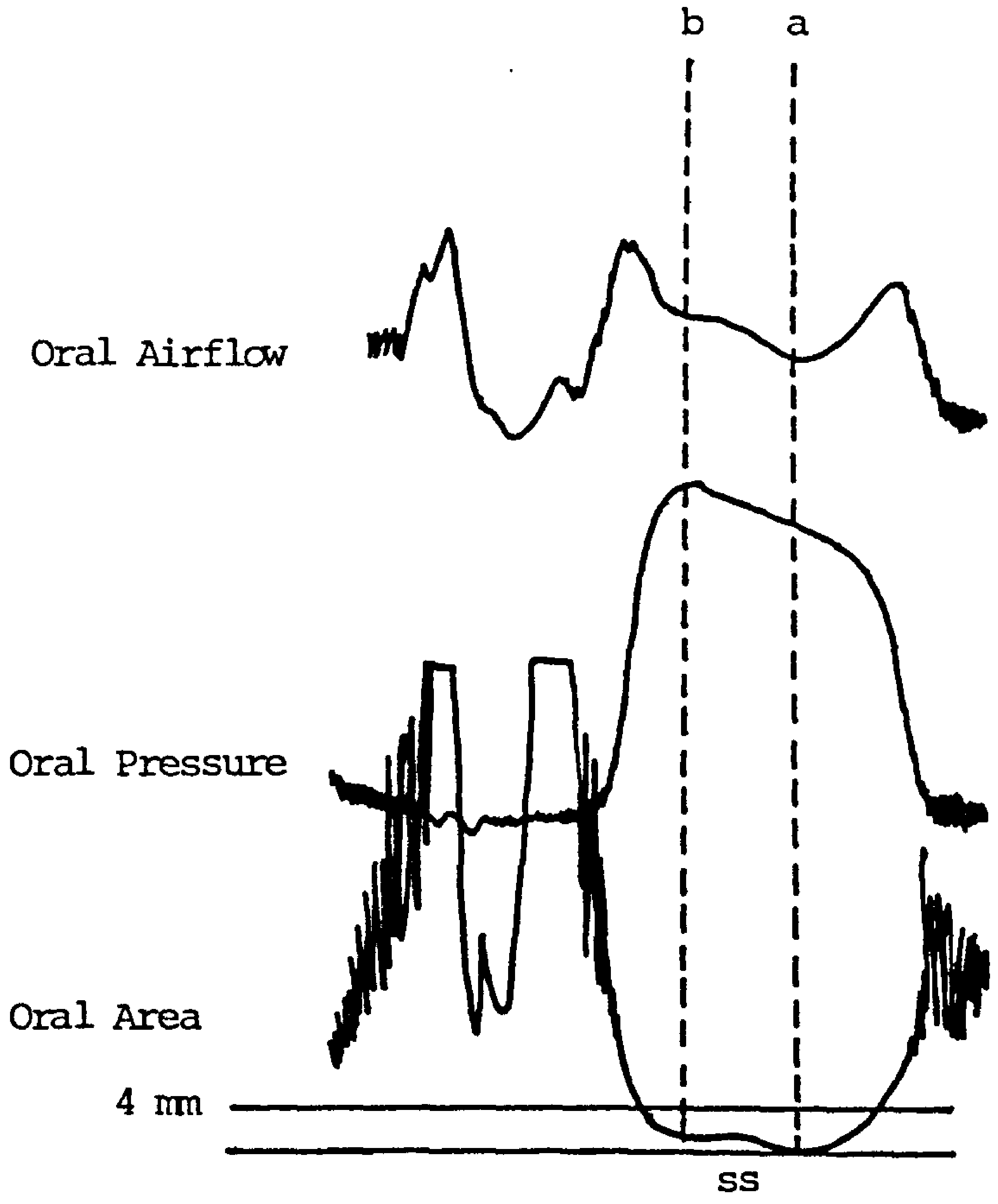


Fig. (6.12) Mingogram of the geminate [ss] in /ħassan/ pronounced by subject AM. a and b are points in time on the 'area' trace.

Table (6.37)

Calculation of Ac at two dips on the 'area' trace. The dips represent two points in time of a and b on that trace for the geminate [ss] in /ħassan/ pronounced by subject AM. The results illustrate that point a has the lower value of Ac.

Item	Points	Airflow (Uo) (cm^3/sec)	Pressure (Po) (cm H_2O)	$Ac = \frac{kUo}{\sqrt{Po}}$ (cm^2)
[ss]	a	305.6	12.1	0.067
	b	416.7	13.8	0.085

threshold value established at 4 mm above the lower point in the dip of the 'area' trace was still operative, and always included the other 'area' dip within the articulatory duration for the fricatives.

As before, two types of intensity level traces were simultaneously recorded on the mingograph together with the other five traces. The low pass trace filtered at 500 Hz was displayed on channel 5, and the high pass trace filtered at 3.9 kHz was displayed on channel 6. Full details of the acoustic traces and the procedures used in our acoustic segmentation and measurements are given in sections 3.2.4 and 6.5.1.

6.11 Statistical Test Used

The results of our perception experiments using synthetic speech and those of our pilot experiment on the aerodynamics and articulation of speech production were subjected to the statistical treatment of the Mann-Whitney U-test; a non-parametric test which can be used with independent-subjects designs. The main reasons behind our choice of the Mann-Whitney test for those experiments were because the U-test is considered as the most powerful non-parametric test for two independent samples which disregards the homogeneity of variance and the normality of the shape of population distributions from which the random samples are drawn. It merely takes account of the rank order of the scores and allows the researcher to test the more general hypothesis that one set of scores tends to be higher, or lower, than another set.

The equivalent parametric test to the Mann-Whitney U-test is the t-test for independent samples which assumes that the population distributions are normal and of the same variance. While the non-parametric tests are, to a certain extent, free from assumptions, the t-test is usually described by statisticians as being extremely robust with respect to violation of these assumptions. Robson (1973) states that "This means there can be considerable deviation from normality and/or homogeneity of variance without the result of the t-test being affected. An exception to this is with the independent-samples design where there are different numbers of scores under the two experimental conditions." [p.72]. However, one should bear in mind that the non-parametric tests do not normally test exactly the same thing as the corresponding parametric tests. The t-test, for instance, tests for a difference in the means of the populations; whereas the corresponding non-parametric test, namely the Mann-Whitney U-test, tests whether or not the populations are identical. (Robson, op.cit., p.107).

In contrast with the parametric tests, the non-parametric tests make very few assumptions about the nature of experimental data. Most non-parametric tests assume only an ordinal level of measurement. Parametric tests, on the other hand, assume that the two samples of data are drawn from populations that have normal distributions and equal variances, i.e. homogeneity of variance. They also assume that the dependent variable has been measured on an interval scale or a ratio scale.

So far, if it comes to the question of which tests to choose; the parametric or the non-parametric tests, statisticians claim that parametric tests are generally more powerful than the corresponding non-parametric tests. Miller (1975), for example, remarks that the power of a test is its ability to detect a significant difference between two sets of scores, and that parametric tests "make use of all the information in the data, whereas the equivalent non-parametric tests simply take account of the rank order of the scores." [p.67]. Roscoe (1975) states that the Mann-Whitney U-test is a non-parametric alternative to the t-test which may be used with ordinal measures or with data that deviate from the normal or in the situation where heterogeneous variances are encountered. He considers the U-test to be almost as powerful as the t-test under common research conditions, but it becomes less powerful if many tie scores are met. It is especially useful with small samples since "violations of the assumptions underlying the t-test are most likely to be detrimental and most likely to go undetected with small samples." [p.230]. Siegel (1956), on the other hand, contends that for a given research, the t-test may be inapplicable for various reasons. He states that the researcher may find that "(a) the assumptions of the t-test are unrealistic for his data, (b) he prefers to avoid making the assumptions and thus to give his conclusions greater generality, or (c) his 'scores' may not be truly numerical and therefore fail to meet the measurement requirement of the t-test." [p.96]. In instances like these, Siegel (loc.cit.) adds, the researcher may choose to analyze his data with one of the non-parametric statistical tests for two independent samples.

Despite all that has been said about parametric and non-parametric tests, researchers are advised always to use a parametric test when the properties of the data allow them to do so, and to settle for the somewhat less powerful non-parametric tests if one or more of the requirements of the parametric tests is violated.

In the light of the above facts, the results of the main experiment have been subjected to the statistical treatment of a 2-tailed t-test for two independent-samples design. This is partly because the t-test is a very powerful test that lends itself to a wide variety of research problems, and partly because our data represent large samples for which the t-test is very effective and, most probably, an appropriate choice. It was felt that 6 tokens were sufficient to justify use of the t-test. In addition, since our data include a variety of independent and dependent variables that might interact intricately, and since the t-test concerns itself with experiments where the measures come from two samples only, the results have been also subjected to the statistical treatment of the analysis of variance (ANOVA for short) which is considered as one of the most general statistical tools as it covers a wide range of possible experimental designs. (Meddis, 1973). ANOVA provides a statistical procedure that is appropriate for use with two or more samples, and in its simplest form, i.e. one-way classification, it is used to test the significance of the differences between the means of a number of different samples.

Generally, ANOVA is a technique for dividing the variation observed in experimental data into different parts, each part assignable to a known source, cause, or factor. (Ferguson, 1959). The basic concept underlying the procedure of ANOVA is that of consistency of scores within the sample. Meddis (op.cit.) states that "If the sample scores show marked fluctuations themselves, then this reduces the significance of small differences between the average scores for the two samples." [p. 2]. He also adds that ANOVA has the specific purpose of dealing with uncertainty produced by using random sampling designs and that it acknowledges the possibility of making errors and attempts, in the significance level statement, to assess just how likely it is that the conclusions one has made are indeed in error. Summarizing, there are four major assumptions that need to be satisfied in applying ANOVA. These assumptions are: independent observations within sets, i.e. random assignment, equal variances within sets, normal distributions population values within sets, and additivity of component contributions to variances.[1]

With regard to our present data, the tables showing the results of ANOVA do not only indicate the p values but

---

[1] For further information concerning the four major assumptions underlying ANOVA, the reader is advised to consult advanced works dealing with statistical tests (e.g. Lindquist, 1940; Ferguson 1959; Scheffé, 1959; Huitson, 1966; Myers, 1966; Roscoe, 1975; Guilford and Fruchter, 1978) and the most useful introductory text by Meddis (1973).

also the F values. The latter values would help the reader discover whether the assumption of homogeneity of variance (i.e. equal variance) was violated by our data. All the results of our statistical analyses were obtained by using the SPSS package available through the main computer terminal at the University Computing Service area.

6.12 Analysis of Results

6.12.1 Method Used for Examining Results

Earlier in this chapter, it was stated that our data included several dependent variables as well as several independent ones. There were, in fact, eight variables of the first type and five of the other type. Our objective was to show what effect, or effects, the independent variables, especially those related to differences in speech rate and to single versus geminate contrasts, have on the dependent ones. Therefore, the entire results of the two subjects who participated in this experiment were first arranged in separate columns and then typed into the computer creating a data file, as shown in tables 6.38 - 6.45. These columns comprise measurements of either the independent or the dependent variables and each column consists of 24 items on each page, thus tokens are of identical conditions, i.e. 6 tokens in each case. The first five, numbered from A to E, are of the independent variables. They are: A for 'speaker', B for 'speech rate', C for 'position of consonant', D for 'type of consonant' and E for 'single versus geminate', in that order.

Since in analyzing data results the computer deals with figures rather than with phrases, the two numbers 1 and 2 were used to signify the two values that each independent variable comprises and to help us in subjecting our data to the statistical treatments. In columns A-E, shown in tables 6.38 - 6.45, number 1 represents subject GG in column A, slow speaking rate in column B, word-initial position in column C, a fricative consonant in column D, and a single consonant in column E. On the other hand, number 2 represents subject AM in column A, fast speaking rate in column B, word-medial position in column C, a plosive consonant in column D, and a geminate consonant in column E.

The other eight columns indicate measurement values pertaining to the eight dependent variables under investigation. In situations where measurements were not possible, the symbol -1 was employed to denote a blank measurement value. Column six shows the measurements of the articulatory timing of the vowel preceding the investigated consonant (T_1), column seven shows the measurements of the articulatory timing between Ac min. of [ħ] and that of [s] in /ħasan/ or [ss] in /ħassan/ (T_4), column eight shows the measurements of the articulatory duration of the consonant (T_2), column nine shows the measurements of the minimum cross-sectional area of word-initial and word-medial [s] and [ss] (Ac), columns ten and eleven show the measurements of oral pressure (Po) and oral airflow (Uo) for the studied consonants, respectively, column twelve shows the measurements of the acoustic duration of the consonant (t_2), and finally column thirteen shows the measurement of the acoustic timing of the vowel preceding that consonant (t_1). Since T_3 does not represent an independent measurement, it was not included in the data file.

In addition to the above eight dependent variables, the computer was also instructed to compute two extra dependent ones at a later stage while creating our SPSS file. The two extra variables were termed 'Arvcratio' and 'Acvcratio' which respectively indicate the ratio obtainable from dividing the articulatory duration of the consonant (T_2) by the articulatory timing of the preceding vowel (T_1), and the ratio obtainable from dividing the acoustic duration of the consonant (t_2) by the acoustic timing of the preceding vowel (t_1).

The SPSS file was created for the purpose of subjecting the results included in the data file to the statistical treatment of the analysis of variance (i.e. ANOVA) and the t-test. This file contained, among other things, a list of the thirteen variables stated in the data file. The thirteen variables were listed exactly in the same order as that shown in the data file, i.e. the list began with the 'speaker' variable and ended with the 't_1' variable. The file also showed the number of cases (i.e. tokens) to be treated in each column.

The SPSS file was also used for t-tests. It also computed the dependent variables of Arvcratio and Acvcratio, mentioned earlier in this section.

For the sake of making things clearer and easier to comprehend both by the investigator and by the reader, the thirteen variables were fully defined and labelled in the SPSS file. The value labels for the five independent variables were also fully described.

6.12.2 The Main Effects of the Independent Variables

In tables 6.46 - 6.50, the main effects of the independent variables on the ten dependent variables are presented. These tables show the p value results of ANOVA indicating whether or not the five independent variables, or combinations of them, have significant effects on the dependent variables. In table 6.46, for example, it is very clear that virtually all independent variables have statistically highly significant effects on the ten dependent variables including the computed ones, namely Arvcratio and Acvcratio, where p value is at < 0.01. The two variables of speaker and consonant type are shown to have statistically highly significant effects on all dependent variables.

Table 6.46 illustrates that differences in speaking rate have statistically non-significant effects on minimum cross-sectional area of the two fricatives [s] and [ss] as well as on oral pressure and oral airflow of the investigated consonants, where p value is at > 0.05. It is also shown that the consonant position in a word has statistically non-significant effects on minimum cross-sectional area of [s] and [ss]. This suggests that there are non-significant differences, in so far as Ac min. is concerned, between the single [s] or the geminate [ss] occurring word-initially, on the one side, and those occurring word-medially, on the other.

The most interesting thing this table shows is that the independent variable of single versus geminate has statistically non-significant effects on the acoustic timing of the preceding vowel (t_1), where p value is at 0.119. This suggests that acoustically there are non-significant differences between the duration of a vowel preceding a single consonant and that of the same vowel preceding a geminate counterpart. These findings are completely consistent with those presented in our acoustic experiment of Chapter Three as well as those presented in our pilot experiment of the present chapter, where it was reported that acoustic durational differences of vowel [a] preceding [s] and [ss] or [d] and [dd], in word-medial positions, were found to be negligible, see sections 3.4 and 6.7.7. These findings add further evidence in support of the view that temporal compensation between the lengthening of word-medial consonants and the preceding vowel does not exist in I.C. Arabic.

In tables 6.47 - 6.50, the effects of the interactions of two, three, four and five independent variables on the dependent ones are shown. Table 6.47, for instance, shows the p values of ANOVA indicating the degree of effect a combination of two independent variables has on the dependent ones. A vertical inspection of this table reveals clearly that almost all interactions of the independent variables have statistically highly significant effects on T_4, i.e., the articulatory timing between Ac min. of [h] and that of [s] in /ħasan/ or [ss] in /ħassan/. T_4 is shown to be less affected by an interaction of the two independent variables of speaker and single versus geminate, where p value is at 0.014 indicating statistically significant effects rather than highly significant ones. At the same time, this table illustrates that the dependent variables representing the articulatory duration of a consonant (T_2), the minimum cross-sectional area of supraglottal constriction (Ac) and the acoustic timing of a preceding vowel (t_1), are shown to be the least affected variables by these interactions of the independent variables.

On the other hand, a horizontal inspection of table 6.47 shows that an interaction of the two independent variables of speaker and speech rate has statistically highly significant effects on T_4 and Po, and statistically significant effects on T_1, T_2 and t_2. Their effects on Ac, Uo, t_1 Arvcratio and Acvcratio are shown to be statistically non-significant. The combination of speech rate and consonant position in a word is shown to have statistically highly significant effects on T_1, T_4, t_1, Arvcratio and Acvcratio, with p value at 0.000; whereas their effects on the other five dependent variables are statistically non-significant, especially on those variables representing aerodynamic features of speech production, viz. Ac, Po and Uo. The results indicate that the effects of the interaction of speech rate and consonant type are statistically highly significant on T_4 and Po, significant on t_2, and non-significant on the other seven dependent variables.

Regarding the effects resulting from the interactions of the independent variable of single versus geminate with the other four independent variables, table 6.47 illustrates that a combination of speaker and single versus geminate has statistically non-significant effects virtually on all dependent variables except on T_4, i.e. the articulatory timing between Ac min. of [ħ] and that of word-medial [s] and [ss], where the effect is shown to be statistically significant, with p value at 0.014. The combination of consonant position and single versus geminate is shown to have statistically non-significant effects on T_2, Ac, Uo and t_1, and statistically significant effects on the other six dependent variables. The interaction of consonant type and single versus geminate seems to have statistically highly significant effects on T_4, Ac, Uo and t_2, and statistically significant effects only on the computed variable of Acvcratio. Their effects are statistically non-significant on the other dependent variables, including T_1 and t_1, i.e. the articulatory and the acoustic timings of the preceding vowel, respectively. In so far as the interaction of rate and single versus geminate is concerned, the results indicate that a combination of the two variables has statistically highly significant effects on five of the dependent variables, namely T_4, T_2, t_2, Arvcratio and Acvcratio; whereas their effects are non-significant on the other five, including T_1 and t_1.

Moreover, table 6.47 shows interactions between other independent variables. For example, the results indicate that a combination of the independent variables of speaker and consonant position has statistically highly significant effects on all dependent variables except T_2 and t_2; while a combination of speaker and consonant type shows statistically highly significant effects on all dependent variables except T_2, t_1 and Arvcratio. Finally, a combination of consonant position and consonant type has statistically highly significant effects on all dependent variables except on Ac and t_2, where the effects are non-significant on the first and only significant on the other.

The results presented in table 6.47 demonstrate very clearly that in combinations where the variable single versus geminate interacts with any of the other four independent variables, there are always statistically non-significant effects on t_1, i.e. the acoustic timing of the vowel preceding the studied word-medial consonants. This is consistent with the results shown in table 6.46.

Table 6.48 shows combinations where three independent variables interact. In this table as well as in tables 6.49 and 6.50, no columns indicating the p values for the two computed dependent variables, i.e. Arvcratio and Acvcratio, are shown. This is because p value results are not possible in three, four and five interactions for the two variables. Nevertheless, the results presented in table 6.48 indicate that the three-way interactions of the independent variables have statistically highly significant effects on T_4, except the tri-combination of the variables of speaker, consonant position and single versus geminate, and that of speaker, consonant type and single versus geminate, where the effects are shown to be statistically significant, with p value at 0.014 in both cases. On the other hand, the results of this table reveal that all three-way interactions have statistically non-significant effects on t_1, which is shown to be less affected by interactions where the two independent variables of rate and single versus geminate coexist with other independent variables. Both T_2 and t_2 are shown to be statistically highly affected by the interaction of speaker, rate and single versus geminate. T_1 seems to be statistically highly affected by the interaction of speaker, consonant position and consonant type, and by the interaction of speaker, consonant type and single versus geminate, suggesting that whenever the two variables speaker and consonant type interact there is statistically high significant effect on T_1, except when they exist together with the independent variable of rate. Oral airflow (Uo) is only affected by the interaction of speaker, consonant position and consonant type, where the effect is shown to be statistically highly significant. This combination has the same effect on Ac. In addition to the statistically highly significant effect of the interaction of speaker, rate and single versus geminate, T_2 is similarly affected by the interaction of consonant position, consonant type and single versus geminate indicating that the independent variable of single versus geminate exerts great effect on the articulatory duration of the consonant.

In table 6.49, the effects of 4-way interactions are presented. These interactions are shown to have statistically highly significant effects on the dependent variable representing the articulatory timing between Ac min. of [ħ] and that of word-medial [s] or [ss] (i.e. T_4) except for the combination of speaker, consonant position, consonant type and single versus geminate where the effect is shown to be only statistically significant, with p value at 0.014. T_1 is shown to be statistically highly affected by the interaction of the independent variables of speaker, speech rate, consonant type and single versus geminate, on the one hand, and by the interaction of the independent variables of speaker, consonant position, consonant type and single versus geminate, on the other. The results of this table illustrate clearly that statistically non-significant effects are shown by the interactions of four independent variables on oral pressure (Po), oral airflow (Uo), acoustic duration of consonant (t_2) and acoustic timing of preceding vowel (t_1). As for the articulatory duration of the consonant (T_2) and the minimum cross-sectional area of supraglottal constriction (Ac), only significant effects are shown by the interaction of speaker, speech rate, consonant position and consonant type on the former, and by the interaction of speaker, speech rate, consonant type and single versus geminate on the latter. The other interactions have non-significant effects on both variables.

Table 6.50 shows combinations where all five independent variables interact. It is obvious that these interactions have statistically significant effects only on T_1 and T_4.

The effects on other dependent variables, namely T_2, Ac, Po, Uo, t_1 and t_2, are shown to be statistically non-significant.

The results of tables 6.46-6.50 show the importance of the two independent variables of speech rate and single versus geminate and indicate their significant effects, either individually or in combination, on the dependent variables associated with the articulatory, aerodynamic, or acoustic aspects of speech production. In order to achieve our objectives in finding out the differences and similarities between utterances spoken with two different speaking rates and in examining contrasts between single and geminate consonants, the data presented in tables 6.38-6.45 were subjected to the statistical treatment of the t-test to see what effect, or effects, each of the above two independent variables has on the ten dependent variables if all other independent variables are kept constant.

6.12.3 The Effect of Speech Rate on the Dependent Variables

Tables 6.51-6.66 show the effects speech rate has on the dependent variables when the other four independent variables remain unchanged. In table 6.51, for example, it is apparent that for subject GG the two different speech rates have statistically highly significant effects on oral pressure (Po) and the acoustic duration of consonant (t_2), suggesting that there are significant differences in Po and in t_2 between word-initial single fricatives produced with slow speaking rate and those produced with fast speaking rate.

For this subject, there is greater Po for fricatives produced with fast speaking rate than for fricatives produced with slow speaking rate. Conversely, t_2 is found to be longer in words spoken with slow speaking rate than in those spoken with fast speaking rate, see lines 1-12 of table 6.38. Table 6.52 shows that for subject AM differences in speech rate have non-significant effects on all dependent variables, indicating that there are statistically non-significant differences, in so far as the five dependent variables are concerned, between word-initial single fricatives used in words spoken with slow speaking rate and those used in words spoken with fast speaking rate, see lines 1-12 of table 6.42.

In tables 6.53 and 6.54, it is quite obvious that for both subjects differences in speech rate have statistically highly significant effects on articulatory and acoustic durations of geminate fricatives pronounced word-initially, suggesting that there are significant durational differences, articulatorily and acoustically, between geminate fricatives produced with two different speaking rates. A geminate fricative used initially in words said with slow speaking rate has been found to be considerably longer than that used in words said with fast speaking rate, see lines 13-24 of tables 6.38 and 6.42. The results of tables 6.53 and 6.54 also show that for the two subjects there are statistically non-significant differences in oral pressure (Po), oral airflow (Uo) and minimum cross-sectional area (Ac) between initial geminate fricatives used in words spoken with slow and fast speaking rates, suggesting that the geminates are not necessarily characterized by having different aerodynamic conditions when produced with two different speech rates.

In tables 6.55-6.58 the effects of speech rate on ten dependent variables are shown. Tables 6.55 and 6.56 indicate similar trends to those presented in tables 6.51 and 6.52. That is, for subject GG significant differences in oral pressure (Po) and the acoustic duration of consonant (t_2) have been found between medial single fricatives pronounced in words said with slow speaking rate and those pronounced in words said with fast speaking rate. The effect of the difference in speech rate is statistically shown to be highly significant on Po and only significant on t_2. The results presented in lines 1-12 of table 6.39 illustrate that there is greater amount of oral pressure in producing a word-medial single fricative spoken with fast speaking rate than that spoken with slow speaking rate. They also illustrate that a word-medial single fricative pronounced with slow speaking rate is acoustically longer than that pronounced with fast speaking rate. In contrast to those results presented in table 6.52, the results of table 6.56 indicate that for subject AM the difference in speech rate does not have significant effects on any of the ten dependent variables, suggesting that for this subject there are non-significant differences between a single fricative used medially in words spoken with slow speaking rate and that used in words spoken with fast speaking rate, see lines 1-12 of table 6.43.

Regarding word-medial geminate fricatives, table 6.57 shows that for subject GG differences in speech rate have significant effects on most of the ten dependent variables. These effects are statistically shown to be either highly significant or only significant, suggesting that there are significant differences between medial geminate fricatives pronounced in words said with slow speaking rate and those pronounced in words said with fast speaking rate, see lines 13-24 of table 6.39. Table 6.58, on the other hand, shows that for subject AM there are significant effects exerted by having different speech rates on only four of the dependent variables, namely T_2, t_1, t_2 and Arvcratio. These effects are statistically shown to be only significant, see lines 13-24 of table 6.43.

The results presented in table 6.57 and 6.58 demonstrate clearly that while subject GG keeps a constant articulatory scheme, subject AM, on the other hand, keeps a constant acoustic scheme, i.e. the vowel and the consonant are reduced in the same proportion acoustically. Subject GG is shown to scale down the articulatory duration and allow the acoustic duration ratio of the vowel and the consonant to vary; whereas subject AM is shown to behave conversely, i.e. he scales down the acoustic duration and allows the articulatory duration ratio of the vowel and the consonant to vary.

In tables 6.59-6.66, the effects of having two different speaking rates in connection with plosive consonants are shown. Tables 6.59 and 6.60 indicate that for the two subjects the articulatory duration of a word-initial single plosive is significantly affected by having two different speech rates. The effect is statistically shown to be highly significant for subject GG and only significant for subject AM. These findings suggest that there are significant differences between the articulatory duration of a single plosive pronounced initially in words spoken with slow speaking rate and that pronounced in words spoken with fast speaking rate, see lines 1-12 of tables 6.40 and 6.44. In as far as t_2 is concerned, the results indicate that for subject GG there are non-significant differences between the acoustic duration of word-initial single plosives produced with slow speaking rate and those produced with fast speaking rate. For subject AM, the situation is different. Table 6.60 shows that for this subject there is a statistically highly significant difference between t_2 of word-initial single plosive consonants. The results indicate that single plosives used word-initially are, acoustically, considerably longer when pronounced with slow speaking rate than those pronounced with fast speaking rate, see lines 1-12 of table 6.44.

Tables 6.61 and 6.62 show the effects of having two different speech rates on word-initial geminate plosives. The two tables indicate clearly that there are significant effects on the articulatory and the acoustic durations of the consonant. For subject GG the effects are statistically shown to be highly significant; whereas they are statistically only significant for subject AM. Articulatorily and acoustically, these results suggest that there are significant durational differences between geminate plosives produced initially in words spoken with slow speaking rate and those produced in words spoken with fast speaking rate. Word-initial geminate plosives have been found to be considerably longer when pronounced with slow speaking rate than those pronounced with fast speaking rate, see lines 13-24 of tables 6.40 and 6.44. At the same time, tables 6.59-6.62 show that oral pressure (Po) is not significantly affected by having two different rates of speech. That is, there are statistically non-significant differences in oral pressure between word-initial plosives, singles or geminates, pronounced in words said with slow speaking rate and those pronounced in words said with fast speaking rate. These results apply to both subjects.

Tables 6.63 and 6.64 show the effects speech rate has on word-medial plosive consonants. Table 6.63 indicates that for subject GG different rates have significant effects on T_1, T_2, Po and t_1. These effects are described as statistically highly significant on the last one. Table 6.64, on the other hand, indicates that for subject AM there are statistically highly significant effects on T_2, Po, t_2 and Arvcratio. The results of the two tables suggest that there are significant differences, in so far as these variables are concerned, between single plosives produced medially in words spoken with slow speaking rate and those produced in words spoken with fast speaking rate, see lines 1-12 of tables 6.41 and 6.45.

Tables 6.65 and 6.66 illustrate the effects of speech rate on word-medial geminate plosives. In table 6.65, it is shown that the articulatory and the acoustic aspects are significantly affected by having two different speaking rates. The effects on the articulatory and the acoustic timing of the preceding vowel (i.e. T_1 and t_1) are shown to be statistically significant; whereas the effects on the articulatory and the acoustic durations of the investigated consonants (i.e. T_2 and t_2) are shown to be statistically highly significant. These effects have their consequences on the p values of the two computed variables, viz. Arvcratio and Acvcratio, which are also shown to be significantly affected. For subject AM, table 6.66 shows that differences in speaking rate have their significant effects only on T_2 and t_2, where the effects are statistically highly significant on the first variable and statistically only significant on the second. Both tables show that there are non-significant differences in oral pressure between geminate plosives produced medially in words spoken with slow speaking rate and those produced in words spoken with fast speaking rate, see lines 13-24 of tables 6.41 and 6.45.

However, the overall results presented in tables 6.63-6.66 yield similar trends to those implied in tables 6.55-6.58. That is, the two subjects are shown to continue keeping two different schemes. So, while subject GG scales down the articulatory duration and allows the acoustic duration of the vowel and the consonant to vary, subject AM does the reverse, i.e. he scales down the acoustic duration and allows the articulatory duration of the vowel and the consonant to vary.

6.12.4 The Effect of Single versus Geminate on the Dependent Variables

Tables 6.67-6.82 show the effects of single versus geminate on the dependent variables when the other four independent variables are kept constant. For example, the p value results presented in tables 6.67 and 6.68 indicate clearly that for both subjects the independent variable of single versus geminate has statistically highly significant effects on T_2 and t_2, suggesting that there are significant differences in the articulatory and acoustic durations of single fricatives and those of their geminate counterparts pronounced with slow speaking rate in word-initial position, see lines 1-6 and 13.18 of tables 6.38 and 6.42. Table 6.67 also shows that for subject GG there are significant differences in oral pressure (Po) between word-initial single fricatives and their geminate cognates when they are pronounced with slow speaking rate. For subject AM, on the other hand, table 6.68 shows that there are significant differences in oral airflow (Uo) and in minimum cross-sectional area (Ac) of word-initial single fricatives and those of their geminate cognates when they are pronounced with slow speaking rate.

The results shown in tables 6.69 and 6.70 indicate very clearly that the independent variable of single versus geminate has significant effects on all dependent variables, except that for subject GG oral pressure seems to be non-significantly affected, where p value is at 0.296. These results show that there are statistically highly significant differences in articulatory and acoustic durations between single fricatives and their geminate partners occurring initially in words spoken with fast speaking rate, see lines 7-12 and 19-24 of tables 6.38-6.42. These findings are completely consistent with those discussed in our foregoing experiments in that single fricatives used word-initially are considerably shorter than their geminate counterparts, see tables 6.20 and 6.29.

In tables 6.71 and 6.72 the effects of single versus geminate on the ten dependent variables are shown when single fricatives or their geminate cognates occur medially in words spoken with slow speaking rate. Both tables indicate that there are statistically highly significant effects exerted by the variable of single versus geminate on T_2 and t_2 as well as on Arvcratio and Acvcratio, suggesting that there are significant differences in articulatory and acoustic durations between single fricatives and their geminate partners produced medially in words spoken with slow speaking rate. The results of the two tables show that whereas for subject GG there are statistically highly significant differences between the articulatory timings of T_1 and T_4 in words having medial single fricatives and in those having their geminate counterparts when produced with slow speaking rate, there are statistically non-significant differences in T_1 and T_4 for subject AM, see lines 1-6 and

13-18 of tables 6.39 and 6.43. The aerodynamic aspects are shown to be non-significantly affected, except that for subject GG it is obvious that there is a statistically significant effect on oral airflow (Uo). These results illustrate clearly that there are non-significant differences in oral pressure (Po) and in minimum cross-sectional area (Ac) between single fricatives and their geminate cognates used medially in words spoken with slow speaking rate.

Both in table 6.71 and in table 6.72, differences in the acoustic timing of the preceding vowel (t_1) are shown to be statistically non-significant, suggesting strongly that there are negligible differences between the duration of a vowel preceding a word-medial single consonant and that preceding a geminate counterpart. These findings are entirely compatible with those reported in our previous experiments, see sections 3.4 and 6.7.7. They give further evidence that there is no temporal compensation between the lengthening of a word-medial consonant and a preceding vowel in I.C. Arabic.

In the case of producing single and geminate fricatives used medially in words spoken with fast speaking rate, tables 6.73 and 6.74 show that the dependent variables of T_2 and t_2 as well as those of Arvcratio and Acvcratio are significantly affected by the independent variable of single versus geminate. The results denote that there are statistically highly significant differences in the articulatory and acoustic durations between a word-medial single fricative and its geminate cognate when both are spoken with fast speaking rate, see lines 7-12 and 19-24 of tables 6.39 and 6.43. They also show that t_1 continues to be non-significantly affected by the single versus geminate variable, thus indicating that there are negligible differences between the acoustic duration of a vowel preceding a single consonant and that preceding a geminate consonant. These findings, in turn, add another support in favour of the fact that temporal compensation between the lengthening of a word-medial consonant and a preceding vowel does not exist in I.C. Arabic.

Tables 6.73 and 6.74 also illustrate that for the two subjects there are non-significant differences in the articulatory timing of a vowel preceding a single consonant and that preceding a geminate cognate when both occur in words spoken with fast speaking rate. Table 6.73 shows that for subject GG there are non-significant differences in oral pressure, oral airflow and minimum cross-sectional area between single fricatives and their geminate partners pronounced medially in words spoken with fast speaking rate. On the other hand, table 6.74 shows that for subject AM there are statistically non-significant differences in oral airflow only, see lines 7-12 and 19-24 of tables 6.39 and 6.43.

In contrast to those results presented in tables 6.57 and 6.58, the results shown in tables 6.73 and 6.74 indicate clearly that the two subjects keep constant articulatory and acoustic schemes in which the vowel and the consonant are reduced in the same proportion.

Tables 6.75 - 6.78 show the effects of single versus geminate on the dependent variables of T_2, Po and t_2. It is apparent that the results of these tables indicate that for both subjects there are significant differences between the articulatory and acoustic durations of single plosives produced initially in words spoken with slow and fast speech rates and their geminate counterparts occurring in the same word-position and produced with the same speech rates. These differences are shown to be statistically highly significant. Nevertheless, table 6.76 illustrates that for subject AM there is a non-significant difference in the amount of oral pressure accompanying the production of a single plosive and that of its geminate partner when both occur initially in words uttered with slow speaking rate, see lines 1-6 and 13-18 of table 6.44.

Tables 6-79-6.82 indicate that the two subjects virtually have identical results. For instance, it is shown that T_2 and t_2 are significantly affected by the independent variable of single versus geminate, suggesting that geminate plosives are, articulatorily and acoustically, considerably longer than their non-geminate cognates when produced medially in words spoken with slow or fast rates of speech. The tables also show that there are negligible differences in the acoustic and articulatory durations of vowels preceding medial single plosives and those preceding their geminate partners no matter whether they occur in words spoken with slow speech rate or spoken with fast speech rate, see measurements presented in tables 6.41 and 6.45. Again, these findings give further evidence to support the idea that there is no temporal compensation between the lengthening of word-medial consonant and a preceding vowel in I.C. Arabic.

Oral pressure (Po) is shown to be non-significantly different in a single plosive and that in a geminate plosive when the two segments occur medially in words spoken with slow or fast speaking rates. However, table 6.82 shows that for subject AM there is a statistically significant difference between oral pressure associated with the production of its geminate cognate when both are used medially in words uttered with fast speech rate. With regard to subject AM, the results of the present experiment indicate that there is a significantly greater amount of oral pressure with a word-medial geminate plosive than with a non-geminate plosive when they are used in words spoken with fast speaking rate, see lines 7-12 and 19-24 of table 6.45.

The conclusions drawn from the results of tables 6.75-6.82 are that fricative and/or plosive geminates are, articulatorily and acoustically, considerably longer than their non-geminate partners whether they are used in words said with slow or fast speaking rates. Differences in the acoustic durations of vowels preceding single consonants and those preceding geminate consonants are found to be negligible. These findings are largely consistent with those of other experiments included in this study, see sections 3.4 and

6.7.7. Oral pressure has been found to vary in relation to the variables of speaker and consonant position. Generally, significant differences in oral pressure have been found between single and geminate consonants when they occur word-initially; whereas non-significant differences have been found when the two segments occur word-medially, irrespective of the speech rate with which the word is uttered. Moreover, the results of table 6.79-6.82 indicate clearly that the two subjects continue keeping constant articulatory and acoustic schemes in which the vowel and the consonant are reduced in the same proportion.

6.12.5 The Interactions Between Single and Geminate Consonants

Speech scientists have always been interested, in so far as the encoding and decoding of speech are concerned, in finding out what things are phonetically common among different speakers and what things are phonetically common across different speech rates. That is, they have been trying to discover cases where there is constancy, i.e. invariance, among things and whether the timing pattern of the linguistic segments stays the same at different speech tempi.

Earlier in this chapter, see section 6.8.1, it has been reported that current speech researches are being increasingly directed towards finding acoustic and articulatory invariance across changes in speech rate. Therefore, the data of our present experiment were subjected to further statistical treatments, in addition to those whose results have been presented in the foregoing tables, to find out whether there are any overlappings between single consonants produced with slow speech rate and their geminate counterparts produced with fast speech rate.

Tables 6.83 - 6.89 show the degree of overlapping between geminate fricatives or plosives used either initially or medially in words spoken with fast speaking rate and their single cognates used in words spoken with slow speaking rate. In tables 6.83 and 6.84, it is clear that for the two subjects there are significant differences in the articulatory and acoustic durations of word-initial and word-medial single fricatives produced with slow speaking rate and their geminate partners produced with fast speaking rate. Although the articulatory and the acoustic durations of the two segments are shown to be statistically highly significant, table 6.83 indicates that for subject AM there is an overlapping in the acoustic duration of a word-initial single fricative and that of its geminate cognate when the former is produced with slow speaking rate and the latter with fast speaking rate. Inspecting the results shown in table 6.42, lines 1-6 and 19-24, it is obvious that there is an interaction in the acoustic duration of the single [s] and that of the geminate [ss] produced word-initially, where the measurement value of 100.8 msec occurs in both segments.

Similarly, tables 6.85 and 6.86 show that there are statistically highly significant differences between the articulatory and the acoustic durations of single plosives used initially or medially in words spoken with slow speaking rate and those of their geminate counterparts used in words spoken with fast speaking rate. The results presented in lines 1-6 and 19-24 of table 6.44, however, show that for subject AM there is an overlapping in the articulatory duration of the single [d] and that of the geminate [dd] produced word-initially, where the measurement value of 85 msec exists in both segments.

The overall results shown in tables 6.87, 6.88 and 6.89 indicate that the interactions noticeable in tables 6.42 and 6.44, in so far as the acoustic and the articulatory durations of the consonant are concerned, are found to be negligible or statistically non-significant. Tables 6.87 - 6.89 show that there are statistically highly significant differences between the acoustic and the articulatory durations of word-initial and word-medial single consonants produced with slow speaking rate and those of their geminate cognates produced with fast speaking rate, suggesting that geminate consonants are very often considerably longer than non-geminate consonants even when they are produced with two contrasting rates of speech.

6.13 Discussion

6.13.1 The Aerodynamic Data

One of the main objectives of speech research is to understand the mechanism underlying the control of the speech-generating system. (Klatt et al., 1968). Several kinds of experimental observations can be made so as to investigate the nature of this process. For instance, aerodynamic events, i.e. those related to airflow and pressure, during speech can provide useful information concerning the activities of various articulatory structures. Measurements of airflow and pressure in speech can give information about the relative timings and speeds of movements of the articulators. (Scully, 1971).

The results obtained from our pilot experiment of this study have shown that geminate consonants are generally characterized by having higher oral pressure than their non-geminate counterparts. For the fricatives, oral airflow has been shown to be greater for the singles than for the geminates, and minimum cross-sectional area has been found to be significantly smaller for a geminate fricative than for a single cognate, see tables 6.4-6.19. These findings have been confirmed in our main experiment for both subjects, where the results have indicated significant differences in oral pressure (Po) between single fricatives or plosives and their geminate partners. Also, the results have indicated significant differences in oral airflow (Uo) and in minimum cross-sectional area (Ac min.) between single fricatives and their geminate cognates. Despite the fact that differences in oral pressure between the geminates and the non-geminates have been found to vary from one speaker to another, in relation to their position in a word, the overall results of the main experiment revealed that geminate consonants are produced with greater oral pressure than the non-geminate consonants and that significant differences in oral pressure have been found between the two types of segments, see table 6.46. Oral airflow has been consistently shown to be greater with single fricatives than with their geminates no matter whether they were articulated with slow or with fast speech rates.

These findings are partly consistent with results reported by Hassan (1981) who examined articulatory and aerodynamic distinctions between a voiceless geminate plosive and its non-geminate counterpart. He observed that the longer occlusion duration of the geminate consonant was accompanied by a higher intraoral pressure though in some cases the geminate and the non-geminate consonants were both stressed. Hassan followed Catford (1977) in suggesting that "this may be indicative of a higher articulatory effort accompanying the act of moving and holding the articulators to maintain a longer occlusion time for the geminate contoid." [p.298]. For Catford (ibid., p.210) a geminate is taken to mean a 'strong' or 'tense' consonant; it involves continuity of articulation, i.e. a prolongation of the articulatory posture, and minor diminution and re-establishment of initiator power in the case of an utterance like [azz-a], or initiator power diminution in the case of an utterance like [az-za]. We do agree with Hassan (loc.cit.) in that it is impossible to confirm, within the limitation of our current investigation, that the articulation of a geminate consonant could be correlated with strong articulatory effort. To achieve this end, one needs to investigate which muscles are responsible for this 'strong' or 'tense' articulation. It may be appropriate here to draw attention to the fact that Arabic is not only characterized by having contrasts between single and geminate consonants, but its phonological system also has oppositions between voiced single and voiced geminate consonants, on the one side, and their voiceless counterparts, on the other.[1] In distinguishing between, let's say, [dd] and [tt], it seems that voicing almost certainly must be maintained throughout the [dd] closure. In order to achieve this long fully voiced closure, the oral pressure must rise more slowly for the geminate [dd] than for the single [d]; probably with voice-maintaining mechanisms.

Speech researchers (e.g. Klatt et al., op.cit.) state that resistance to airflow during speech production comes primarily from two sources : the glottis and constrictions in the supraglottal vocal tract. Scully (op.cit., p.192) points out that resistance depends upon the length of the constriction as well as upon the cross-sectional area. She states that since the airflow is turbulent, resistance is then a function of airflow.

Our measurements of the oral airflow for fricatives show that single consonants are accompanied by slightly greater oral airflow than geminate consonants, see tables 6.8, 6.9 and 6.38-6.45. Considering these results, we can postulate that since both [s] and [ss] are voiceless consonants, requiring very little obstruction of the airflow by the vocal folds such as that observed during the articulation of voiced consonants, what then really causes the consider-

---

[1] See Chapter One of this study for full details.

able differences in oral airflow between the two homogeneous segments is the degree of constriction in the supraglottal vocal tract. The geminate [ss] has been found to be produced with a significantly smaller minimum cross-sectional area than the single [s]. This strongly suggests that there is closer approximation between the tongue tip or body and the alveolar ridge during the articulation of [ss] than during the articulation of [s], thus resulting in greater oral pressure and lower oral airflow for the geminate and lesser oral pressure and higher oral airflow for the non-geminate. These findings are highly consistent with those of the direct palatography experiment, which suggested strongly that the geminate [ss] is produced with firmer tongue-palate contact than the single [s]. These results, once again, lend support to the notion of 'mechanical pressure' originally advanced by McGlone and Proffit (1972) implying that there is greater force exerted by the tongue against the roof of the mouth during the articulation of the geminate consonants than that during the articulation of the non-geminate ones, see section 5.5 for more details.

Although our current study is not principally designed to investigate differences between voiced single or geminate consonants and their voiceless partners, our data present evidence that there is higher oral pressure during the production of single and geminate fricatives than during the production of single and geminate plosives. The reason may lie in the role played by the glottis as being a major source of obstruction while producing the two voiced plosive consonants, namely [d] and [dd], where a part of the air-pressure is used for the vibrating vocal folds. In this respect, Isshiki and Ringel (1964) state that "it is felt that the closure of the glottis, a necessary requisite for voicing, increases the resistance within the vocal tract, and therefore reduces the flow rate." [p.241]. Slis (1970, p.195) explains that the driving force for the vibration of the vocal folds is supplied by the air-stream through the glottis and that if the intraoral pressure rises, the pressure drop over the glottis will decrease, thus making the air-stream through the glottis decrease. He argues that with voiceless consonants this happens rather quickly giving no time to keep the vocal folds in vibration; whereas with voiced consonants the pressure drop remains sufficiently high to keep the vocal folds in vibration. His results indicated that the glottis was more open during voiceless consonants than during voiced consonants. On the same point, Arkebauer et al. (1967) state that resistance of the speech air-stream at the level of the vocal folds probably results in a pressure drop across the glottis during voiced consonant productions, and that the glottis offers little resistance to airflow during the production of voiceless sounds. They assume that the respiratory system produces higher subglottal pressures during voiced than during voiceless consonant productions so that intraoral pressures produced during voiced consonants may approach the magnitudes of those produced during voiceless consonants. Further, they state that "although it is not known whether or not maximum subglottal air pressures differ for voiced and voiceless consonants, it appears obvious that the pressure drop across the glottis in the former type of speech sound is so great that the variable of voicing is a primary determinate of intraoral air pressures associated with consonant productions." [pp.204-205].

It is commonly agreed (Black, 1950; Isshiki and Ringel, 1964; Hixon, 1966; Arkebauer et al., 1967; Scully; 1971) that voiceless consonants are accompanied by substantially higher airflow and by greater oral pressure than voiced consonants. Perkell (1969), and Kent and Moll (1969) have demonstrated that voiced plosives are produced with larger supraglottal volumes than their voiceless cognates. Moreover, Kent and Moll's data revealed that pharyngeal expansion occurs simultaneously with the depression of the hyoid bone, which is probably an active process. They concluded that the increase in supraglottal volume is at least a partial explanation for instraoral pressure differences among cognates.

The data from Warren and Hall's study (1973) suggested that pressure differences among voiced-voiceless cognate pairs are related to the length of the voicing interval. The two researchers noted that the most significant difference in pressure was observed when voicing occurred throughout the period of pressure rise and the smallest difference was obtained when voicing was not apparent. These findings support the contentions of a number of investigators (e.g. Malécot, 1955; Warren and Wood, 1969) that vocal fold activity produces the largest pressure drop across the vocal tract. According to them, this does not necessarily mean that vocal fold activity is the primary determinant of intraoral pressure.[1] Warren and Wood (op.cit.) have suggested that respiratory effort and intraoral volume are important factors which may influence consonantal pressures. They reported that respiratory volumes (i.e. airflow rates) are consistently larger for voiceless sounds and proposed that intraoral air volumes may differ as a result of this.

Our results representing the aerodynamic conditions of speech production are comparable with those presented by other investigators. They are, for instance, in good agreement with Black's findings. Black (op.cit.) reported that his data suggested that greater magnitudes of intraoral air pressure are associated with voiceless consonants than with their voiced cognates and that greater pressures are associated with fricative sounds than with plosive sounds. On the other hand, they are in disagreement with findings reported by Malécot (1955) and Subtelny et al. (1966). Malécot's research does not support Black's conclusion as well as ours that higher peak intraoral air pressures are associated with fricatives than with plosives. Malécot stated that peak magnitudes of intraoral air pressure are greater during the production of voiceless plosives than during voiceless fricatives; while the trend is reversed for voiced plosives and fricatives. Similarly, Subtelny etal. (op.cit.) found that a higher amplitude of oral pressure was

---

[1] Cf. Arkebauer et al., 1967, p.205 .

recorded during voiceless plosive articulation than during voiceless fricative articulation.

Other studies on the voice source (e.g. van den Berg, 1956; Ladefoged and McKinney, 1963; Scully, 1969) have shown that increases in the level of voice are accompanied by increases in subglottal air pressure and that air pressures generated within the vocal tract increase as the level of speech is raised. These studies have posited that generally higher voice intensities are accompanied by greater rates of airflow through the vocal tract at successively higher levels of speech production, whether the speech signal is a vocal tone or a turbulent noise. Scully's investigation presents evidence that, in general, volume airflow rate through the glottis and volume airflow rate through the tongue constriction are not equal. She argues that they differ partly because of changes in the supraglottal cavity volume and partly because of changes inside the supraglottal cavity, which means that pressure-airflow experiments do not give direct information about the airflow at the vocal folds. She adds that it is not easy to measure the subglottal pressure and, therefore, it is difficult to make statements about the glottal resistance.

With regard to data obtained from fricative consonants, our experiments present results compatible with those reported by Hixon (1966). Hixon's data suggested that both greater intraoral air pressures and oral airflow rates were associated with higher noise levels. He states that "for voiceless fricatives, where the laryngeal valve is assumed to be relatively open, the greater rates of airflow into the oral cavity at higher speech intensities would result in greater air pressure magnitudes intraorally providing the area of the maximum oral constriction does not enlarge at a rate proportional to or greater than the rate of airflow." [p.178]. Hassan (1981) has observed that volume airflow rate and intraoral pressure are significantly higher for voiceless fricative and plosive consonants than for their voiced counterparts, and that the minimum cross-sectional area of constriction is significantly smaller for a voiced fricative than for a voiceless partner. His data also showed that peak intraoral pressure is generally greater with the plosive consonants than with the fricative consonants.

Aerodynamically speaking, what our present investigation has revealed is that intraoral air pressures and volume airflow rates are consistently higher for voiceless consonants, than for voiced ones, irrespective of differences in speech rates as well as in speakers. On the other hand, and probably more importantly, the results of the main experiment have quite obviously indicated that oral pressure, oral airflow and minimum cross-sectional area of supraglottal constriction do not change statistically significantly with changes in rate of speech when other factors are held constant. That is, single or geminate consonants produced with fast speech rate show statistically insignificant differences in oral pressure, oral airflow and constriction area from single or geminate consonants produced with slow rate of speech. See tables 6.38 - 6.45 for measurements pertaining to these three variables, and also see table 6.46 for statistical evaluation. Therefore, since the measured aerodynamic and physiological dimensions, in so far as oral pressure, oral airflow and constriction area are concerned, do not change with changes in speech rate, it appears quite logical to assume that during increased speed of single versus geminate articulations the manner in which the examined consonants, viz. [s] ~ [ss] and [d] ~ [dd], are articulated does not change significantly.

Although these findings provide no information as to the precision of the vocal tract shape required for the articulation of single versus geminate consonants, the data obtained strongly suggest that at the two speaking rates used in the main experiment of this study, the area of the minimum oral constriction does not change significantly. When it is considered that neither oral pressure nor oral airflow does change with changing speech rates, it then seems plausible to assume that the shape of the supraglottal constriction also may not have changed at different rates of speech.

These findings appear to be in total agreement with those reported by Hixon (op.cit.) whose data provided evidence that air pressure, airflow and constriction area were not affected by speaking rate. At the same time, it seems reasonable to adopt Hixon's interpretation in suggesting that his findings might have important implications for the acoustic theory of speech production as elaborated by Stevens and House (1963). Based on acoustic data, these researchers suggested that articulatory positioning for vowels tends to approach an ideal target configuration more closely when the consonantal environment is a plosive than when it is a fricative, in spite of the fact that the vowel durations for plosive consonantal environments are shorter than those for fricative environments. They stated that "such a finding is not unexpected if it is considered that the articulatory structures can execute displacements to and from the complete closures such as those appropriate for stop consonants with considerable speed, whereas fricative environments require that structures be accelerated from and decelerated to configurations with constrictions of precisely controlled size and shape." [p.126]. They believed that it is probable that the movement to and from such fricative configurations should be less rapid and, therefore, the extent to which the vowel configuration falls short of the ideal target configuration is greater for fricative consonantal environments.

In the light of the above argument and in so far as our data are concerned, it is evident that the main experiment presents evidence that the articulation of the geminate consonants or that of the non-geminate consonants does not require relatively precisely controlled size and shape constrictions within the vocal tract when that segment is produced with either slow or fast rates of speech. This finding, however, does not agree with results reported by Stevens and House (op.cit.) and Lindblom (1963) which indicate that changes in the dynamics of speech physiology are to be expected when speech rate is varied. That is, when the data of the main experiment are considered, the lack of systematic trends for significant changes in intraoral pressure, oral airflow rate and the inferred minimum cross-sectional area associated with increasing speech rate during the articulation of single or geminate consonants makes the interpretation of these results difficult. We do, therefore, agree with the contention proposed by Hixon (op.cit.) in that parameters such as air pressure, airflow and oral constriction to that airflow should be taken into account for a thorough understanding of the aerodynamics of consonant articulation.

Finally, one of the most controversial points that the current investigation has addressed is that our data bear testimony to the fact that gemination in Arabic does not involve a rearticulation (i.e. double articulation) of the same consonant. The obtained traces, representing the aerodynamic, articulatory and acoustic conditions of single versus geminate production, were almost without exception steady and uninterrupted for either the friction noise of a geminate fricative or the occlusion time of a geminate plosive. Such findings are comparable with those presented by Hassan (op.cit.) who stated that "neither from the acoustic nor from the myodynamic and aerodynamic points of view is there any indication that gemination involves a rearticulation of the same consonant. They all show steady and uninterrupted traces for the occlusion time of the geminate plosive." [p.396]. Hegedüs (1959) also found that Hungarian geminates are characterized by the non-existence of rearticulation in their production.

Our present findings are also compatible with hypotheses suggested by other investigators (e.g. Blanc, 1952; Nasr, 1960; Erwin, 1963, 1969; Cowell, 1964; Al-Ani, 1970; Al-Ani and May, 1973) in that a geminate consonant in Arabic is articulated as one long indivisible consonant. They are, on the other hand, in discrepancy with findings reported in investigations conducted on languages other than Arabic (e.g. Stetson, 1951; Lehiste, 1960, 1966, 1970; Delattre, 1971). These scholars presented evidence that gemination does involve double articulation of the same consonant in the languages they studied. Lehiste et al. (1973) have found evidence for rearticulation in the production of geminate consonants in Estonian, but their data provided no evidence for rearticulation in the production of English consonant sequences at word boundaries.[1]

6.13.2 The Acoustic and the Articulatory Data

The measurements of the acoustic and articulatory durations which we have cited in the tables of the pilot and the main experiments of this study indicate very clearly that geminate consonants occurring word-initially or word-medially are considerably longer than their non-geminate partners occurring in the same word positions. The significant durational differences between single and geminate consonants

---

[1] For more references and further information, see Chapters One and Two of this study.

have been found to be applicable to consonants produced in words spoken with different rates of speech by two native speakers of I.C. Arabic.

If we look back at tables 6.29-6.32 for acoustic durations and at tables 6.20-6.23 for articulatory durations, it is seen that a geminate fricative is virtually twice as long as its single cognate, and a geminate plosive is more than three times the duration of its single counterpart. These findings are highly consistent with results presented in our experiment of Chapter Three. They are also highly consistent with findings reported by other investigators (e.g. Han, 1965; Delattre, 1971; Lehiste et al., 1973; McKay, 1980, Hassan, 1981; Balasubramanian, 1982).

As reviewed in section 2.3, Han (ibid.) conducted a spectrographic investigation and found that 'long' (geminate) consonants in Japanese are much more than twice the duration of their 'short' (single) counterparts. For example, he compared the word /hakada/ with /hakkada/ and the word /kitari/ with /kittari/, and discovered that "the closure of /k/ in /hakada/ lasts 11.5 centiseconds while the closure of long /kk/ in /hakkada/ lasts 31.5 centiseconds. In the case of /kitari/ ~ /kittari/, the closure of the short /t/ is 8.5 centiseconds while that of the long /tt/ is 25.0 centiseconds." [p.71]. Delattre (ibid.) and Lehiste et al. (ibid.) studied the phenomenon of gemination in a variety of languages and found out that in these languages a geminate consonant is significantly longer than its single cognate, and that consonant duration is regarded as a major factor in the linguistic functioning of gemination. McKay (op.cit.) came to the same conclusion when he studied single and geminate plosive consonants in Rembarrnga. He stated that differences in closure duration between geminate and single plosives found in his study for velars and bilabials were comparable with those found by Lehiste et al. between single and short geminate Estonian plosives.

Hassan (op.cit.) investigated vowel duration in I.C. Arabic and presented results indicating that the acoustic and articulatory durations of a word-medial geminate plosive are significantly longer than those of its single partner. The two words /fat'taat/ and /fa'taat/ were studied. The results showed that the acoustic duration of the geminate "is even more than twice of that for the non-geminate." [p.399]. Our present findings are also in agreement with results reported by Balasubramanian (op.cit.) who obtained kymographic tracings revealing that the duration of the 'double' (geminate) nasals and laterals in Indian Tamil is two and a half times that of their single counterparts.

Also, if we look back at tables 6.33 and 6.34, it is clear that the durational differences between the vowel [a] preceding a single consonant and that preceding a geminate consonant are acoustically non-significant. The mean values of [a] are shown to be insignificantly longer before the geminates than those before the singles. From table 6.27, it is also clear that there are articulatorily non-significant differences between the vowel [a] preceding a single plosive and that preceding a geminate cognate. These results are, likewise, consistent with those presented in Chapter Three of this study, see tables 3.15 and 3.16 for /ħasan/ versus /ħassan/ and tables 3.23 and 3.24 for /badal/ versus /baddal/. They strongly suggest that in I.C. Arabic vowels occurring in stressed positions maintain their original length when they precede word-medial geminate consonants. They are, once again, compatible with results reported by Delattre (op.cit.) and McKay (op.cit.). They confirm that the duration of the preceding vowel is considered as a negligible factor when distinguishing a geminate consonant from its single counterpart.

Nonetheless, these findings are in contradiction with results presented by other researchers who provided evidence indicating that vowels preceding single consonants are predominantly longer than those preceding geminate consonants. Hassan (op.cit.), for instance, contended that the shorter acoustic duration of the preceding vowel could be ascribed to the advance in closure time of the following geminate plosive consonant. He agreed with Slis (1967) in suggesting that the advanced closure time of the geminate consonant is part of articulatory timing planned at a higher linguistic level in which the speaker deliberately executes an earlier occlusion onset and maintains a longer occlusion time for the geminate for phonological reasons. Hassan believed that the shorter acoustic duration of the preceding vowel and the longer acoustic duration of the following geminate resulted from a similar articulatory programme for the supraglottal articulation. He claimed that these acoustic durations could be part of a pattern serving in the perceptual realization of gemination as a phonological phenomenon. Therefore, he agrees with neither Delattre's nor with our results in that duration is considered as a major, and probably the only, perceptual cue in distinguishing a geminate from a single consonant. He assumed that the shorter acoustic duration of the preceding vowel "also contributes to the perceptual realization of the following long consonants as geminate consonants which are acoustically shorter and preceded by a relatively longer vocoid." [p.399].

However, the perceptual findings of our experiments on synthetic speech, see section 4.6, showed that listeners could easily distinguish geminate consonants from their non-geminate partners when only the consonant duration was varied while keeping the duration of the preceding vowel unchanged. These findings were in good agreement with Delattre's contentions in that his perceptual tests confirmed that consonant duration is a major cue for the perception of gemination and that the duration of the preceding vowel is not a factor in the perception of consonant gemination. His results indicated that vowels were not significantly shorter before a geminate than before a single consonant.

Our main experiment does not only show that a geminate consonant is significantly longer than its single cognate when the two segments are produced initially or medially in words uttered with slow speaking rate or in words uttered with fast speaking rate, but it also shows that a geminate consonant produced in words spoken with fast rate of speech is predominantly longer than a non-geminate partner produced in words spoken with slow speech rate, see tables 6.83 - 6.89. Evidently, this suggests that there is no chance of overlapping, in so far as duration is concerned, between a geminate consonant and a single cognate and that a geminate is always considerably longer than a non-geminate under any circumstances. Following Lehiste (1970, pp.12-13)[1] in supposing that the range of DL's duration is usually from 10 to 40 msec in sequences ranging from 30 to 300 msec, it is quite obvious that the durational differences between single and geminate consonants, as shown in Figs. 6.13 -6.16, are well above DLs. These figures suggest strongly that a listener can distinguish a single consonant from a geminate cognate when both segments are produced with slow or with fast speech rates. They also suggest that a listener would distinguish all single consonants from all geminate ones.

However, at this stage of our discussion of the relation between the duration of a consonant and a preceding vowel, we believe it is appropriate to refer to the concept of temporal compensation which we have already mentioned in section 3.3. The idea of temporal compensation has been extensively discussed by a number of researchers to show its phonetic significance in different languages (e.g. Kozhevnikov and Chistovich, 1965; Lindblom, 1968; Lindblom and Rapp, 1973;

---

1. See section 4.6.2 for more information.

Chen, 1970; Lehiste, 1972; Ohala, 1975; Klatt, 1976; Port et al., 1980). These studies have been primarily designed to investigate the extent to which temporal compensation takes place within spoken utterances. That is, the extent or domain of interaction defining the size of the phonological unit that is temporally programmed at some level of the speech formation process. (Wright, 1974). They have also aimed at determining which linguistic unit, i.e. the syllable, the word, or a larger unit, is the basic unit of speech timing. For instance, Lehiste (1970, 1971) reported results which showed that temporal compensation involved all segments within the represented monosyllabic and bisyllabic CCVC words that she investigated. She found compensation taking place between most adjacent phonemes in both types of words. Specifically, she always found it occurring between a vowel and a following consonant in monosyllabic words. Wright (op.cit.) comments saying that this latter finding suggests that "there is a stronger temporal bond or connection between a vowel and following consonant than between a consonant and following vowel, indicating that the VC syllable, rather than the CV syllable, might be the basic unit of timing for English utterances." [p.1258]. In another study designed by Allen (1970), evidence has been presented suggesting that the CV syllable is not the basic unit of timing in English. On the other hand, Shockey et al. (1971) found no evidence supporting either the VC or CV syllable as the basic timing unit for English utterances. They found significant negative correlations between the durations of certain adjacent segments within a word when it was spoken in a larger context.

The present experiments do not provide evidence in support of the idea of temporal compensation. The measurements presented in the tables, particularly those of /ħasan/ versus /ħassan/ and /badal/ versus /baddal/, do not show any evidence of compensation between the lengthening of word-medial consonant and a preceding vowel. Durational differences between vowels preceding geminate consonants and those preceding non-geminate consonants are found to be negligible. Our findings are, to a certain extent, compatible with those reported by Port et al. (op.cit.)[1] who stated that their results give minimal evidence of temporal compensation in their examined syllables in Arabic and that there is no evidence that vowel lengthening for voicing of a following consonant is related to the shortening of the consonant itself. Moreover, neither on the acoustic nor on the articulatory level is there in our data any indication of the existence of temporal compensatory adjustment between the preceding vowel and the following geminate/non-geminate consonant. Our results of Chapter Three and those of the present chapter indicate very clearly that the overall acoustic and articulatory timings of the sequences /ħas-/ versus /ħass-/ and /bad-/ versus /badd-/ as well as those of the whole words of /ħasan/ versus /ħassan/ and /badal/ versus /baddal/ are significantly different. Evidently, this signi-

---

[1] See section 2.4 for more details.

fies that temporal compensation does not exist between the following geminate/non-geminate consonant and the preceding vowel in I.C. Arabic, in so far as our data are concerned.

These findings may be compared with those reported by Hassan (op.cit., p.397) who assumed that the shorter acoustic duration of the preceding vowel might be compensated for by the duration of the following geminate consonant was ruled out by the fact that the overall timings of the first syllable as well as the whole word showed very significant differences. He believes that this implies that there is no indication of a temporal compensatory adjustment between the preceding vowel and the following geminate and/or non-geminate consonant neither on the syllable nor on the word levels. Hassan's findings showed similar trends to those obtained in our current chapter as well as those obtained in Chapter Three of this study. That is, the acoustic and the articulatory results do not show any evidence of compensation between the lengthening of word-medial consonant and a preceding vowel, thus suggesting that durational differences between vowels preceding geminate consonants and those preceding non-geminate consonants are negligible.

With regard to the effect of speech rate on the acoustic and articulatory durations of consonants and their preceding vowels, our data have demonstrated that geminate consonants used initially or medially in words spoken with slow speaking rate are significantly longer than those used in words spoken with fast speaking rate. Single consonants have also been shown to be longer when produced with slow speech rate than when produced with fast speech rate.

Although the results presented in table 6.46 indicate that speech rate has statistically highly significant effects on the acoustic and the articulatory timings of the preceding vowel, suggesting that the acoustic timing (t_1) and the articulatory timing (T_1) of the preceding vowel are significantly longer before consonants articulated in words said with slow speaking rate than before consonants articulated in words said with fast speaking rate, the results presented in tables 6.55 - 6.58 show quite obviously that vowels preceding word-medial fricatives, geminates or non-geminates, have insignificant acoustic and articulatory durational differences when produced with slow and fast rates of speech. Table 6.58, however, shows that for subject AM the acoustic timing of the preceding vowel (t_1) is significantly longer before a geminate fricative produced in words spoken with slow speaking rate than before a geminate fricative produced in words spoken with fast speaking rate. On the other hand, the results shown in tables 6.63 - 6.66 indicate that only for subject GG the acoustic and the articulatory timings of the preceding vowel are significantly longer before single and geminate plosives produced in words said with slow speech rate than those before single and geminate plosives produced in words said with fast speech rate. A close inspection of the acoustic and the articulatory measurements presented in tables 6.38 - 6.45 will reveal the fact that long linguistic segments produced with slow speech rate are liable to major reductions in their durations when they are produced with fast rate of speech. Segments of short duration are less affected by differences in speech rate.

In section 6.8.1, we have stated that investigators have been trying to find out what happens, acoustically and articulatorily, when a talker alters his rate of speech. It was argued that both the segment durations measured from the acoustic record of an utterance and the extent of articulatory movements decrease proportionately with an increase in speaking rate. However, some researchers state that an examination of individual segments indicates that some reduce more than others, and that voiced consonants reduce the least with an increase in speech rate. For instance, Kozhevnikov and Chistovich (1965) and Klatt (1976) have documented that vowels shorten more than consonants when speech rate increases.

It is widely known that changes in speaking rate exert a complex influence on the durational patterns of a sentence. When speakers slow down, a large fraction of the extra duration goes into pauses. (Goldman-Eisler, 1968). Increases in speaking rate are, on the other hand, accompanied by phonological and phonetic simplifications as well as differential shortening of vowels and consonants. Gay et al. (1974) contend that there is a complex reorganization of the motor commands to the articulators such that consonantal gestures are strengthened as speaking rate is increased, but the motor commands for vowels are not enhanced.

Our present data have shown the same tendency as that suggested by Gaitenby (op.cit.) in that not only do vowels shorten at fast rates of speech but also consonants. Nevertheless, Gaitenby also states that the time differences between fast and slow speakers in speech are greatest in vowels in stressed or pre-pause positions. According to her, there are numerous, and frequently simultaneous, possible factors that influence word duration; some of which are physiological, others are linguistic, and still others are specific to the individual talker and the occasion. Our results indicate that the duration of single consonants suffered less reduction than the duration of geminate consonants when produced at fast speech rate. It has been shown that virtually one third of the original duration of a geminate consonant, used either initially or medially in words spoken with slow rate of speech, was lost when that geminate was articulated in words spoken with fast speaking rate, see tables 6.38 - 6.45. We can, then, postulate that the longer the sound segment is the greater the reduction it suffers when produced at rapid rates of speech. Since geminate consonants are longer than single consonants; therefore they are liable to be greatly shortened at fast speaking rates than their single counterparts.

In contrast to the findings reported by Kozhevnikov and Chistovich (op.cit.) our data showed that consonants as well as vowels shorten when speech rate is increased. Gay (1978) found that for stressed vowels, an increase in speaking rate is accompanied primarily by a decrease in duration. That is, the vowel portion is most affected during fast speech. We have found that the vowel [a] preceding the geminate/non-geminate consonants loses a relatively small amount of its original duration when produced with fast rate of speech. This account predicts that, in connection with our studied data, consonants shorten more than vowels when speech rate increases. This finding is inconsistent with the widely accepted fact, originally suggested by Lindblom (1963), that during faster speech a vowel tends to change colour towards a neutral quality, i.e. towards the schwa. Lindblom's original model posits that this neutralization is a consequence of the shorter duration of the vowel and is caused by a temporal overlap of motor commands to the articulators. That is, the articulators fail to reach, or undershoot, their targets because the next set of motor commands deflects them to the following target before the first target is reached. Therefore, in producing a vowel the target is reached and the gesture is completed if the speaking rate is slow, but if it is fast, the movement then is cut off before reaching the final target. (Gay, 1968).

Lindblom's contentions were later discussed by Gay et al. (1974) who argued that whereas vowel production involves a movement toward a spatial target, the production of most consonants involves a movement towards constrictive or occlusal targets, thus suggesting that the concept of undershoot itself cannot be easily applied to consonant production. Their findings indicated converse results to those suggested by Lindblom. They stated that the decrease in the activity level of the genioglossus muscle found by them shows that undershoot is not a consequence of an overlap in the timing of commands to the muscle, and that "the decrease in muscle activity for the tongue during faster speech may reflect only the decrease in overall displacement of the tongue (and not any changes in its speed of movement)." [p.62]. Elsewhere, Gay (1978, p.223) found that generally electromyographic activity associated with tongue body movements during vowel production decreased during fast speech, while activity associated with both labial and alveolar plosive consonant production increased with an increase in speaking rate.

6.13.3 The Concept of Invariance in Relation to Our Data

One of the main objectives of speech research has been to characterize the defining properties of speech sounds that occur in natural languages, and to determine how the listener extracts these properties in the process of speech perception. (Mack and Blumstein, 1983). Recently, a number of scholars (e.g. Cole and Scott, 1974; Blumstein and Stevens, 1980, 1981; Stevens and Blumstein, 1981) have advanced a theory of acoustic invariance which suggests that invariant acoustic properties can be derived directly from the acoustic signal, and that these properties correspond to the phonetic dimensions occurring in natural language. This theory has been elaborated in most detail in connection with the place of articulation for plosive consonants in a series of investigations employing both synthetic and real speech. For instance, Stevens and Blumstein (1978) analyzed a series of synthetic continua representing the phonetic categories [b, d, g] in the context of the three vowels [a, i, u] in terms of the spectral characteristics of the stimuli signifying each phonetic category. The results showed that the spectrum obtained by sampling the first 20-odd msec at consonantal release seems to exhibit invariant properties for each place of articulation for plosive consonants independent of vowel context. (Blumstein and Stevens, 1980, p.649).

Another endeavour has been made by Blumstein and Stevens (1979) to provide a more quantitative measure of the degree to which invariant patterns were derivable for voiced and voiceless consonants across vowel contexts. The results showed that the spectrum at onset does not seem to produce an invariant gross shape for place of articulation independent of the vowel context. Both the burst and the initial part of the transition were shown to contribute to this invariance and they together form a single integrated acoustic property. Following earlier studies, Mack and Blumstein (1983) have stated that "if one focused on the notion of integrated properties associated with the stop release burst and transitions over the initial 25 msec, acoustic invariance could be derived for place of articulation in stop consonants." [p.1739].

Works that dealt with distinction between plosive and approximant consonants have employed synthetic speech to investigate the effects of manipulation of particular acoustic parameters on the perception of the [b-w] contrast. For example, studies carried out by O'Connor et al. (1957), Cooper (1976) and Schwab et al. (1981) have included various parameters such as rate, extent and duration of the formant transitions. The results obtained showed that duration and extent seem to be critical perceptual dimensions. However, other studies suggest that plosives and approximants may differ with respect to a large number of acoustic dimensions, including not only transition duration, rate and extent, but also presence or absence of a burst, presence or absence of low-amplitude periodicity at the onset of the signal, and frequency and amplitude of individual formants at onset. (Fant, 1960; Lehiste and Peterson, 1961; Pickett, 1980). It is assumed that any one of these attributes, or a combination of them, may well contribute to an invariant property which distinguishes plosive consonants from approximant consonants.

More recently, Mack and Blumstein's research (1983) has concentrated on another phonetic contrast to determine the extent to which an invariant acoustic property or properties can be derived to characterize the distinction between a plosive consonant and an approximant. The results of their three experiments indicated that transition durations and formant frequencies often differed considerably in these two classes of sounds, i.e. the plosive-approximant contrast. The results also indicated that larger changes in amplitude energy were associated with the plosive release than with the approximant release across vowel contexts and speakers. In their view, these changes seem to provide an invariant property characterizing the plosive-approximant contrast.

Mack and Blumstein believe that in the theory of acoustic invariance there are implicit claims about the perception of speech. They state that such a theory holds that the perceptual system responds distinctively to the invariant properties derived from the acoustic signal, and that it is the presence of acoustic invariance which structures the phonological system for the speaker-hearer. If we accept this argument, it would then be expected that the perceptual system shown in Chapter Four of this study responds in a distinctive manner to relative durational differences between single and geminate consonants similar to those derived from the production data in the present experiments.

If we consider the findings observed in the whole experiments of the present study, and particularly those of this chapter, we may suggest that durational differences between geminate and non-geminate consonants in I.C. Arabic are of the same order of magnitude as those reported in other studies (e.g. Obrecht, 1965; Delattre, 1971; Lehiste et al. 1973; McKay, 1980). Our findings indicated that absolute duration is a likely, but probably not the only, candidate for an invariant property to distinguish between single and geminate consonants in I.C. Arabic. This is partly because the duration measures for the geminate and the non-geminate consonants do not vary as a function of vowel context, and partly because the perceptual system is reportedly non-sensitive to vowel duration in making a decision about the geminate or the non-geminate identification. It is clear from the acoustic and the articulatory measures presented in this Chapter that duration closures associated with single and geminate plosives are different. They are also different in association with the duration of single and geminate fricatives. The geminates are always considerably longer than their non-geminate partners whether they are used word-initially or word-medially, pronounced by one and the same speaker or by different speakers, produced in words spoken in isolation or in words spoken in a carrier sentence, articulated with slow or with fast speech rates. Therefore, we may exclude suprasegmental factors such as those of speech rate, consonant position, consonant type and speakers from being likely candidates for distinguishing between single and geminate consonants in I.C. Arabic.

Tables of Data

Explanation of Column Headings

A = Speaker (1=GG, 2=AM)·

B = Rate (1 = "slow", 2 = "fast")*

C = Position (1 = initial, 2 = medial)

D = Type (1 = fricative, 2 = plosive)

E = Single/geminate (1 = single, 2 = geminate)

T1= Articulatory timing of preceding vowel

T4= Articulatory timing between Ac minimum of [ħ] and that of word-medial [s] or [ss]

T2 = Articulatory duration of consonant

Ac = Minimum cross-sectional area of supraglottal constriction.

Po = Oral pressure

Uo = Volume airflow rate

Acdurc (t2) = Acoustic duration of consonant

Acdurv (t1) = Acoustic timing of preceding vowel

---

* N.B. "slow" is taken to mean normal speaking rate, as opposed to "fast".

A	B	C	D	E	T1	T4	T2	AC	PO	UO	ACDURC	ACDURV
1	1	1	1	1	-1	-1	62	0.039	16.9	209.6	93.0	-1
1	1	1	1	1	-1	-1	70	0.065	17.5	349.4	93.0	-1
1	1	1	1	1	-1	-1	54	0.081	15.6	419.3	93.0	-1
1	1	1	1	1	-1	-1	62	0.078	17.5	419.3	100.8	-1
1	1	1	1	1	-1	-1	54	0.133	15.0	663.9	85.3	-1
1	1	1	1	1	-1	-1	70	0.079	16.9	419.3	93.0	-1
1	2	1	1	1	-1	-1	47	0.083	18.1	454.2	77.5	-1
1	2	1	1	1	-1	-1	62	0.083	18.1	454.2	85.3	-1
1	2	1	1	1	-1	-1	54	0.075	18.8	419.3	85.3	-1
1	2	1	1	1	-1	-1	62	0.078	20.0	454.2	85.3	-1
1	2	1	1	1	-1	-1	54	0.087	18.8	489.2	77.5	-1
1	2	1	1	1	-1	-1	54	0.078	20.0	454.2	85.3	-1
1	1	1	1	2	-1	-1	217	0.063	22.5	349.4	240.3	-1
1	1	1	1	2	-1	-1	163	0.056	18.1	314.5	178.3	-1
1	1	1	1	2	-1	-1	163	0.062	18.1	349.4	186.1	-1
1	1	1	1	2	-1	-1	163	0.075	18.1	419.3	193.8	-1
1	1	1	1	2	-1	-1	155	0.063	18.1	349.4	162.8	-1
1	1	1	1	2	-1	-1	147	0.070	18.1	384.3	178.3	-1
1	2	1	1	2	-1	-1	85	0.065	17.5	349.4	108.5	-1
1	2	1	1	2	-1	-1	124	0.080	18.8	454.2	139.5	-1
1	2	1	1	2	-1	-1	93	0.071	20.0	419.3	116.3	-1
1	2	1	1	2	-1	-1	124	0.065	20.6	384.3	147.3	-1
1	2	1	1	2	-1	-1	109	0.070	21.3	419.3	124.0	-1
1	2	1	1	2	-1	-1	116	0.072	20.0	419.3	139.5	-1

TABLE 6.38

A	B	C	D	E	T1	T4	T2	AC	PO	UO	ACDURC	ACDURV
1	1	2	1	1	93.0	124.0	54	0.063	17.5	349.4	85.3	69.8
1	1	2	1	1	85.3	108.5	54	0.062	18.8	349.4	93.0	69.8
1	1	2	1	1	93.0	116.3	47	0.071	17.5	384.3	77.5	77.5
1	1	2	1	1	93.0	116.3	47	0.083	17.5	454.2	77.5	69.8
1	1	2	1	1	93.0	124.0	54	0.099	16.3	524.1	85.3	62.0
1	1	2	1	1	93.0	124.0	54	0.075	18.1	419.3	77.4	62.0
1	2	2	1	1	93.0	116.3	47	0.073	18.8	419.3	62.0	69.8
1	2	2	1	1	77.5	108.5	54	0.066	20.0	384.3	77.5	62.0
1	2	2	1	1	77.5	108.5	47	0.072	20.0	419.3	69.8	69.8
1	2	2	1	1	93.0	124.0	54	0.062	22.5	384.3	77.5	62.0
1	2	2	1	1	93.0	116.3	47	0.051	21.9	314.5	77.5	69.8
1	2	2	1	1	85.3	116.3	47	0.078	20.0	489.2	69.8	62.0
1	1	2	1	2	124.0	224.8	171	0.048	19.4	279.5	193.8	85.3
1	1	2	1	2	100.8	209.3	155	0.072	17.5	384.3	162.8	85.3
1	1	2	1	2	131.8	217.1	140	0.062	18.8	349.4	178.3	69.8
1	1	2	1	2	131.8	201.6	124	0.052	17.5	279.5	170.5	69.8
1	1	2	1	2	116.3	217.1	155	0.062	18.8	349.4	178.3	69.8
1	1	2	1	2	124.0	217.1	132	0.065	17.5	349.4	170.5	62.0
1	2	2	1	2	93.0	124.0	116	0.078	17.5	419.3	131.8	69.8
1	2	2	1	2	85.3	131.8	109	0.065	20.6	384.3	131.8	62.0
1	2	2	1	2	77.5	124.0	101	0.058	21.3	349.4	116.3	62.0
1	2	2	1	2	93.0	178.3	109	0.061	24.4	384.3	131.8	77.5
1	2	2	1	2	85.3	124.0	109	0.054	24.4	349.4	124.0	69.8
1	2	2	1	2	100.8	131.8	101	0.077	24.4	489.2	116.3	69.8

TABLE 6.39

A	B	C	D	E	T1	T4	T2	AC	PO	UO	ACDURC	ACDURV
1	1	1	2	1	-1	-1	47	0	12.5	0	31.0	-1
1	1	1	2	1	-1	-1	47	0	13.1	0	31.0	-1
1	1	1	2	1	-1	-1	47	0	13.8	0	38.8	-1
1	1	1	2	1	-1	-1	47	0	13.1	0	38.8	-1
1	1	1	2	1	-1	-1	39	0	12.5	0	38.8	-1
1	1	1	2	1	-1	-1	47	0	11.9	0	31.0	-1
1	2	1	2	1	-1	-1	39	0	12.5	0	38.8	-1
1	2	1	2	1	-1	-1	39	0	13.1	0	31.0	-1
1	2	1	2	1	-1	-1	31	0	12.5	0	31.0	-1
1	2	1	2	1	-1	-1	31	0	14.4	0	31.0	-1
1	2	1	2	1	-1	-1	39	0	14.4	0	31.0	-1
1	2	1	2	1	-1	-1	39	0	14.4	0	38.8	-1
1	1	1	2	2	-1	-1	140	0	15.6	0	131.8	-1
1	1	1	2	2	-1	-1	163	0	18.1	0	147.3	-1
1	1	1	2	2	-1	-1	147	0	15.6	0	131.8	-1
1	1	1	2	2	-1	-1	116	0	15.6	0	108.5	-1
1	1	1	2	2	-1	-1	140	0	16.3	0	124.0	-1
1	1	1	2	2	-1	-1	140	0	16.9	0	131.8	-1
1	2	1	2	2	-1	-1	93	0	15.6	0	93.0	-1
1	2	1	2	2	-1	-1	93	0	15.6	0	85.3	-1
1	2	1	2	2	-1	-1	93	0	16.3	0	85.3	-1
1	2	1	2	2	-1	-1	109	0	17.5	0	100.8	-1
1	2	1	2	2	-1	-1	109	0	15.0	0	93.0	-1
1	2	1	2	2	-1	-1	116	0	18.8	0	100.8	-1

TABLE 6.40

A	B	C	D	E	T1	T4	T2	AC	PO	UO	ACDURC	ACDURV
1	1	2	2	1	85.3	-1	39	0	14.4	0	31.0	85.3
1	1	2	2	1	85.3	-1	39	0	12.5	0	31.0	93.0
1	1	2	2	1	77.5	-1	47	0	13.8	0	38.8	85.3
1	1	2	2	1	77.5	-1	39	0	13.8	0	38.8	85.3
1	1	2	2	1	85.3	-1	39	0	14.4	0	31.0	85.3
1	1	2	2	1	85.3	-1	39	0	13.8	0	31.0	85.3
1	2	2	2	1	69.8	-1	39	0	14.4	0	31.0	77.5
1	2	2	2	1	85.3	-1	31	0	15.0	0	31.0	77.5
1	2	2	2	1	69.8	-1	31	0	13.8	0	31.0	62.0
1	2	2	2	1	69.8	-1	31	0	15.6	0	31.0	69.8
1	2	2	2	1	69.8	-1	31	0	16.9	0	31.0	69.8
1	2	2	2	1	77.5	-1	39	0	16.3	0	38.8	69.8
1	1	2	2	2	85.3	-1	155	0	11.9	0	139.5	93.0
1	1	2	2	2	85.3	-1	163	0	16.3	0	147.3	100.8
1	1	2	2	2	85.3	-1	163	0	15.6	0	147.3	85.3
1	1	2	2	2	77.5	-1	147	0	15.6	0	139.5	85.3
1	1	2	2	2	77.5	-1	155	0	16.3	0	139.5	85.3
1	1	2	2	2	77.5	-1	140	0	15.6	0	131.8	77.5
1	2	2	2	2	77.5	-1	85	0	13.8	0	77.5	85.3
1	2	2	2	2	77.5	-1	109	0	16.3	0	100.8	69.8
1	2	2	2	2	69.8	-1	93	0	16.3	0	93.0	69.8
1	2	2	2	2	77.5	-1	109	0	16.3	0	93.0	85.3
1	2	2	2	2	77.5	-1	101	0	15.0	0	100.8	77.5
1	2	2	2	2	77.5	-1	109	0	16.3	0	93.0	77.5

TABLE 6.41

A	B	C	D	E	T1	T4	T2	AC	PO	UO	ACDURC	ACDURV
2	1	1	1	1	-1	-1	70	0.045	10.0	180.6	112.4	-1
2	1	1	1	1	-1	-1	62	0.056	8.3	208.3	100.8	-1
2	1	1	1	1	-1	-1	78	0.059	9.2	222.2	108.5	-1
2	1	1	1	1	-1	-1	54	0.038	9.6	152.8	116.3	-1
2	1	1	1	1	-1	-1	62	0.051	8.8	194.4	85.3	-1
2	1	1	1	1	-1	-1	62	0.030	10.0	125.0	100.8	-1
2	2	1	1	1	-1	-1	54	0.046	7.9	166.7	100.8	-1
2	2	1	1	1	-1	-1	62	0.056	9.6	222.2	96.9	-1
2	2	1	1	1	-1	-1	62	0.043	9.2	166.7	100.8	-1
2	2	1	1	1	-1	-1	62	0.046	9.2	180.6	81.4	-1
2	2	1	1	1	-1	-1	47	0.068	7.9	250.0	93.0	-1
2	2	1	1	1	-1	-1	70	0.037	9.6	150.0	93.0	-1
2	1	1	1	2	-1	-1	163	0.025	15.8	111.1	162.8	-1
2	1	1	1	2	-1	-1	163	0.014	12.1	55.6	186.0	-1
2	1	1	1	2	-1	-1	178	0.042	10.0	166.7	186.0	-1
2	1	1	1	2	-1	-1	171	0.020	10.0	83.3	178.3	-1
2	1	1	1	2	-1	-1	163	0.027	10.8	111.1	174.4	-1
2	1	1	1	2	-1	-1	163	0.042	10.0	166.7	155.0	-1
2	2	1	1	2	-1	-1	147	0.038	10.4	152.8	147.3	-1
2	2	1	1	2	-1	-1	155	0.038	10.4	138.9	155.0	-1
2	2	1	1	2	-1	-1	132	0.014	10.8	55.6	139.5	-1
2	2	1	1	2	-1	-1	132	0.040	10.4	166.7	131.8	-1
2	2	1	1	2	-1	-1	132	0.035	9.2	138.9	100.8	-1
2	2	1	1	2	-1	-1	140	0.040	10.4	166.7	131.8	-1

TABLE 6.42

A	B	C	D	E	T1	T4	T2	AC	PO	UO	ACDURC	ACDURV
2	1	2	1	1	85.3	104.7	47	0.060	15.4	305.6	62.0	69.8
2	1	2	1	1	108.5	135.7	54	0.063	13.8	305.6	77.5	100.8
2	1	2	1	1	62.0	89.1	70	0.038	11.7	166.7	108.5	54.3
2	1	2	1	1	77.5	104.7	62	0.042	16.3	222.2	96.9	62.0
2	1	2	1	1	77.5	96.9	62	0.044	18.8	250.0	77.5	77.5
2	1	2	1	1	93.0	112.4	62	0.041	13.3	194.4	85.3	100.8
2	2	2	1	1	62.0	96.9	70	0.050	11.7	222.2	69.8	65.9
2	2	2	1	1	93.0	131.8	62	0.099	12.1	444.4	65.9	89.1
2	2	2	1	1	62.0	93.0	47	0.065	12.9	305.6	73.6	62.0
2	2	2	1	1	77.5	93.0	39	0.072	12.9	333.3	62.0	69.8
2	2	2	1	1	93.0	124.0	54	0.041	12.9	194.4	93.0	77.5
2	2	2	1	1	73.6	85.3	54	0.073	12.5	333.3	69.8	62.0
2	1	2	1	2	77.5	116.3	140	0.046	17.5	250.0	155.5	77.5
2	1	2	1	2	85.3	135.7	124	0.048	15.8	250.0	139.5	81.4
2	1	2	1	2	93.0	217.1	194	0.067	13.8	305.6	193.8	73.6
2	1	2	1	2	89.1	135.7	140	0.081	13.8	388.9	170.5	77.5
2	1	2	1	2	93.0	135.7	163	0.039	14.6	194.4	189.9	77.5
2	1	2	1	2	77.5	124.0	147	0.045	14.6	222.2	166.7	69.8
2	2	2	1	2	85.3	127.9	109	0.052	13.8	250.0	124.0	73.6
2	2	2	1	2	69.8	127.9	132	0.030	13.8	138.9	155.0	69.8
2	2	2	1	2	77.5	116.3	101	0.046	13.8	222.2	139.5	62.0
2	2	2	1	2	85.3	127.9	116	0.041	14.6	208.3	147.3	69.8
2	2	2	1	2	85.3	155.0	116	0.049	13.8	236.1	143.4	69.8
2	2	2	1	2	93.0	127.9	124	0.051	13.8	250.0	139.5	73.6

TABLE 6.43

A	B	C	D	E	T1	T4	T2	AC	PO	UO	ACDURC	ACDURV
2	1	1	2	1	-1	-1	70	0	10.4	0	62.0	-1
2	1	1	2	1	-1	-1	54	0	9.6	0	62.0	-1
2	1	1	2	1	-1	-1	62	0	10.4	0	65.9	-1
2	1	1	2	1	-1	-1	62	0	8.3	0	69.8	-1
2	1	1	2	1	-1	-1	62	0	10.0	0	62.0	-1
2	1	1	2	1	-1	-1	85	0	8.8	0	69.8	-1
2	2	1	2	1	-1	-1	39	0	7.1	0	46.5	-1
2	2	1	2	1	-1	-1	62	0	9.2	0	42.6	-1
2	2	1	2	1	-1	-1	54	0	9.6	0	50.4	-1
2	2	1	2	1	-1	-1	47	0	7.1	0	54.3	-1
2	2	1	2	1	-1	-1	47	0	9.2	0	54.3	-1
2	2	1	2	1	-1	-1	47	0	7.9	0	54.3	-1
2	1	1	2	2	-1	-1	147	0	11.7	0	147.3	-1
2	1	1	2	2	-1	-1	140	0	7.5	0	131.8	-1
2	1	1	2	2	-1	-1	155	0	12.1	0	162.8	-1
2	1	1	2	2	-1	-1	155	0	11.3	0	155.0	-1
2	1	1	2	2	-1	-1	163	0	10.8	0	151.2	-1
2	1	1	2	2	-1	-1	155	0	10.0	0	147.3	-1
2	2	1	2	2	-1	-1	140	0	9.6	0	147.3	-1
2	2	1	2	2	-1	-1	124	0	10.4	0	116.3	-1
2	2	1	2	2	-1	-1	109	0	11.7	0	108.5	-1
2	2	1	2	2	-1	-1	85	0	9.6	0	89.1	-1
2	2	1	2	2	-1	-1	132	0	10.0	0	124.0	-1
2	2	1	2	2	-1	-1	140	0	10.4	0	139.5	-1

TABLE 6.44

A	B	C	D	E	T1	T4	T2	AC	PO	UO	ACDURC	ACDURV
2	1	2	2	1	89.1	-1	54	0	12.9	0	50.4	96.9
2	1	2	2	1	77.5	-1	47	5	12.9	0	46.5	85.3
2	1	2	2	1	69.8	-1	54	0	12.5	0	50.4	77.5
2	1	2	2	1	85.3	-1	62	0	$12.9	0	54.3	93.0
2	1	2	2	1	77.5	-1	54	0	10.8	0	46.5	77.5
2	1	2	2	1	85.3	-1	54	0	9.2	0	46.5	100.8
2	2	2	2	1	77.5	-1	47	0	9.6	0	42.6	77.5
2	2	2	2	1	77.5	-1	47	0	8.8	0	46.5	93.0
2	2	2	2	1	77.5	-1	39	0	6.3	0	42.6	77.5
2	2	2	2	1	69.8	-1	39	0	6.7	0	31.0	93.0
2	2	2	2	1	81.4	-1	39	0	8.8	0	38.8	85.3
2	2	2	2	1	77.5	-1	39	0	7.9	0	34.9	77.5
2	1	2	2	2	100.8	-1	171	0	10.8	0	170.5	100.8
2	1	2	2	2	93.0	-1	140	0	12.9	0	135.7	93.0
2	1	2	2	2	85.3	-1	147	0	9.6	0	147.3	85.3
2	1	2	2	2	89.1	-1	140	0	9.2	0	135.7	96.9
2	1	2	2	2	93.0	-1	155	0	10.8	0	155.0	93.0
2	1	2	2	2	77.5	-1	140	0	9.6	0	139.5	96.9
2	2	2	2	2	89.1	-1	124	0	10.0	0	131.8	93.0
2	2	2	2	2	77.5	-1	124	0	9.2	0	131.8	93.0
2	2	2	2	2	85.3	-1	124	0	10.4	0	131.8	85.3
2	2	2	2	2	85.3	-1	124	0	10.4	0	124.0	93.0
2	2	2	2	2	77.5	-1	124	0	10.0	0	127.9	77.5
2	2	2	2	2	77.5	-1	132	0	9.6	0	131.8	77.5

TABLE 6.45
=========

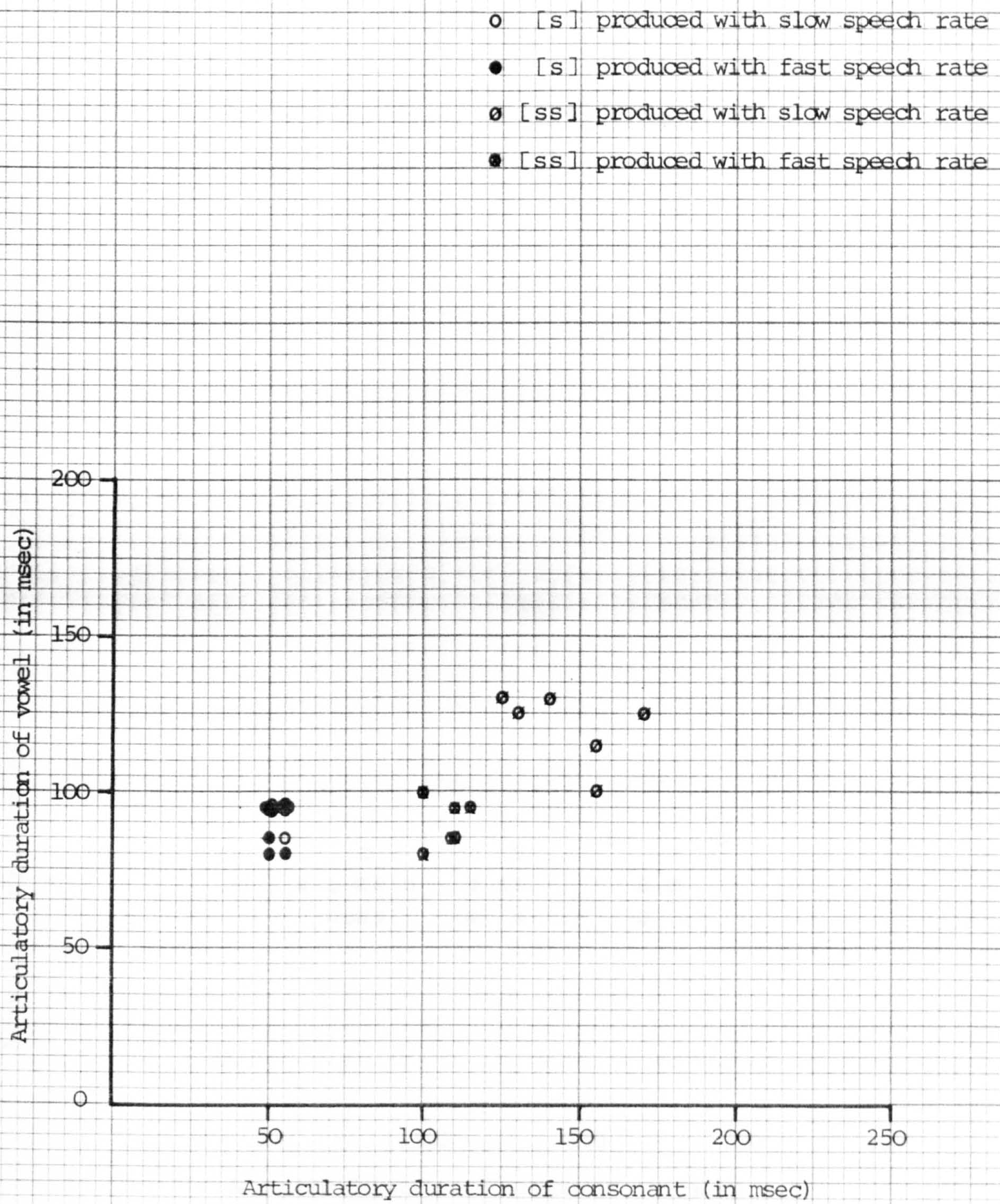


Fig. (6.13) Graph showing articulatory durations for single and geminate fricative consonants (Subject : GG)

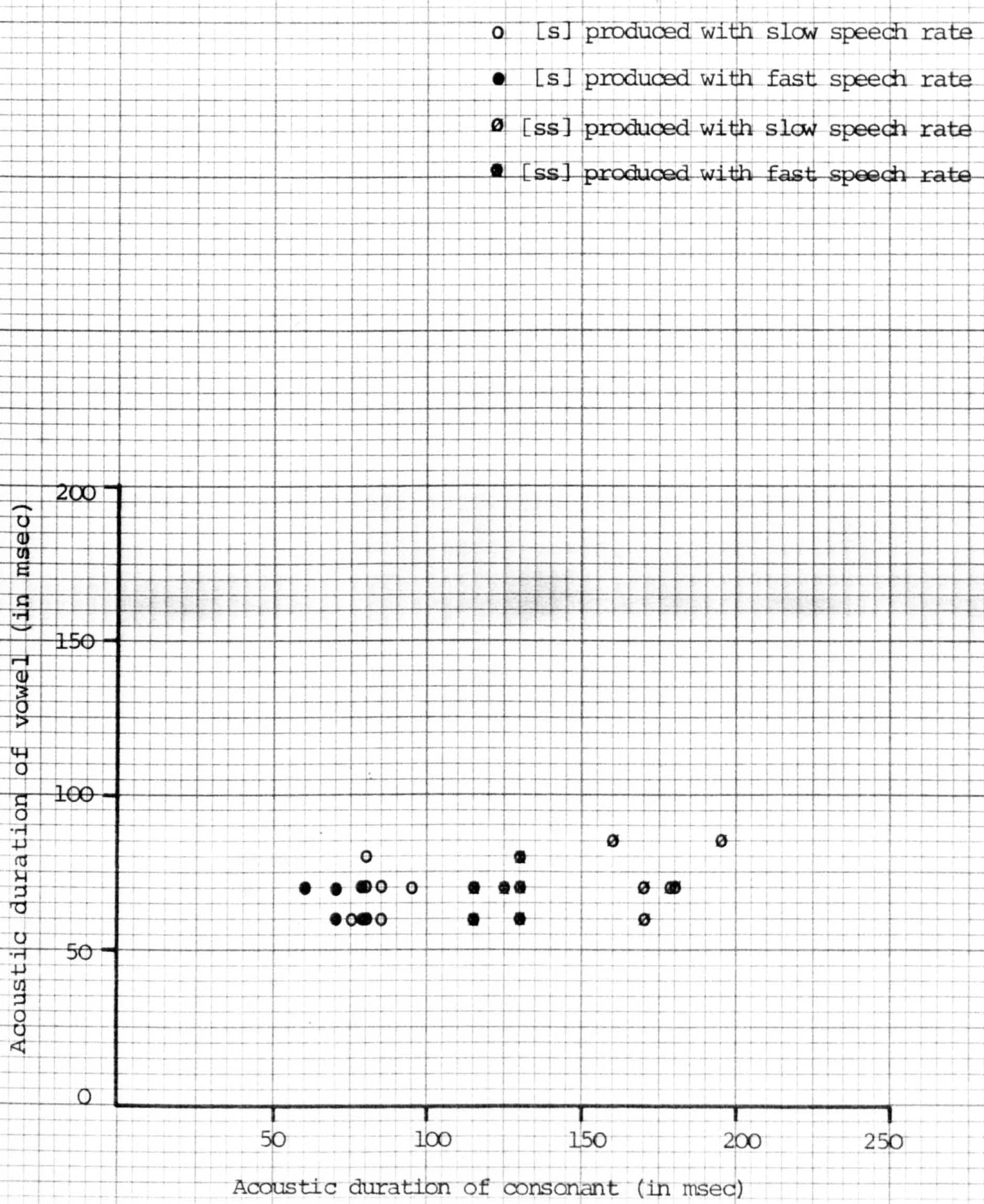


Fig. (6.14) Graph showing acoustic durations for single and geminate fricative consonants (Subject : GG)

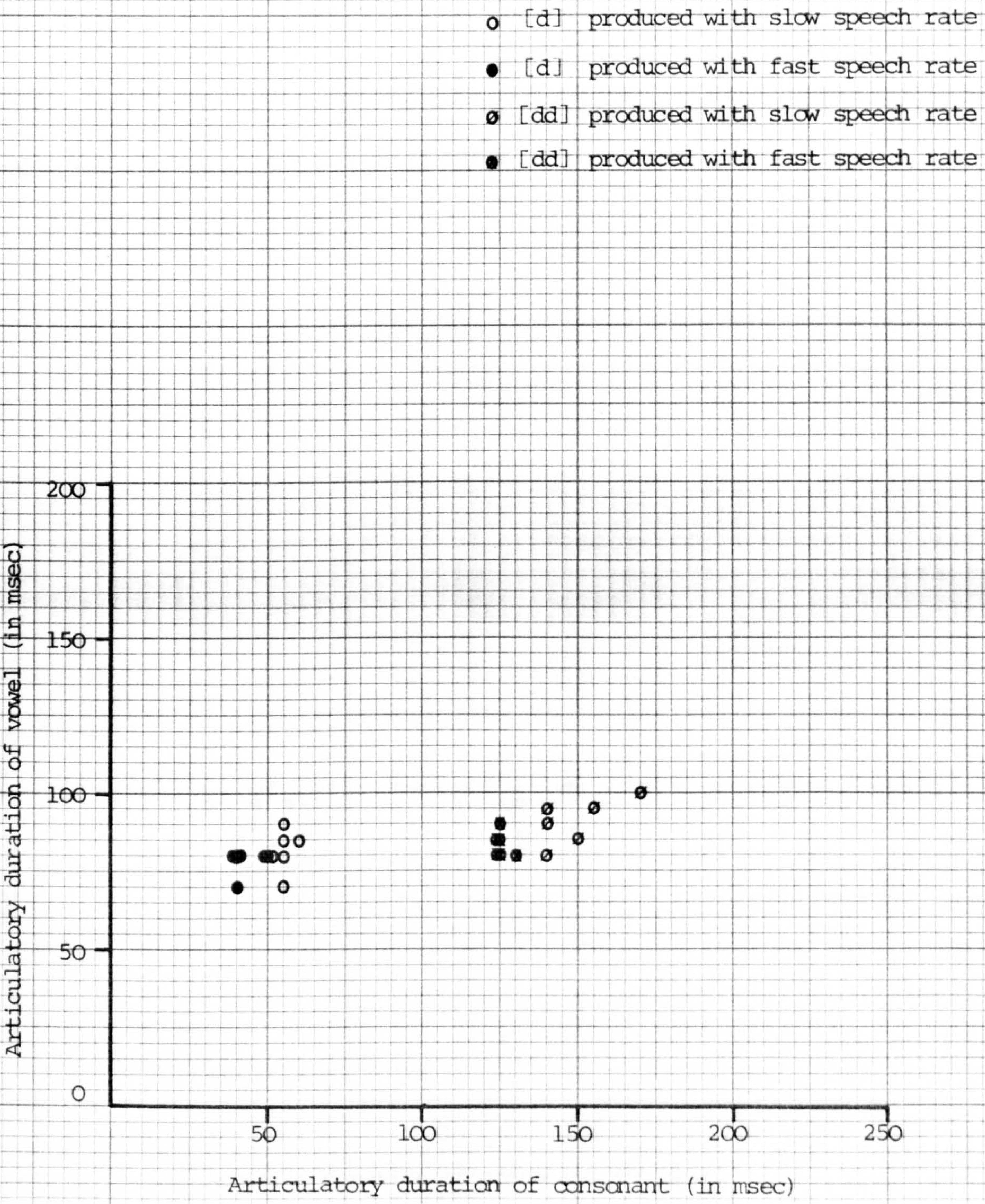


Fig. (6.15) Graph showing articulatory durations for single and geminate plosive consonants (Subject : AM)

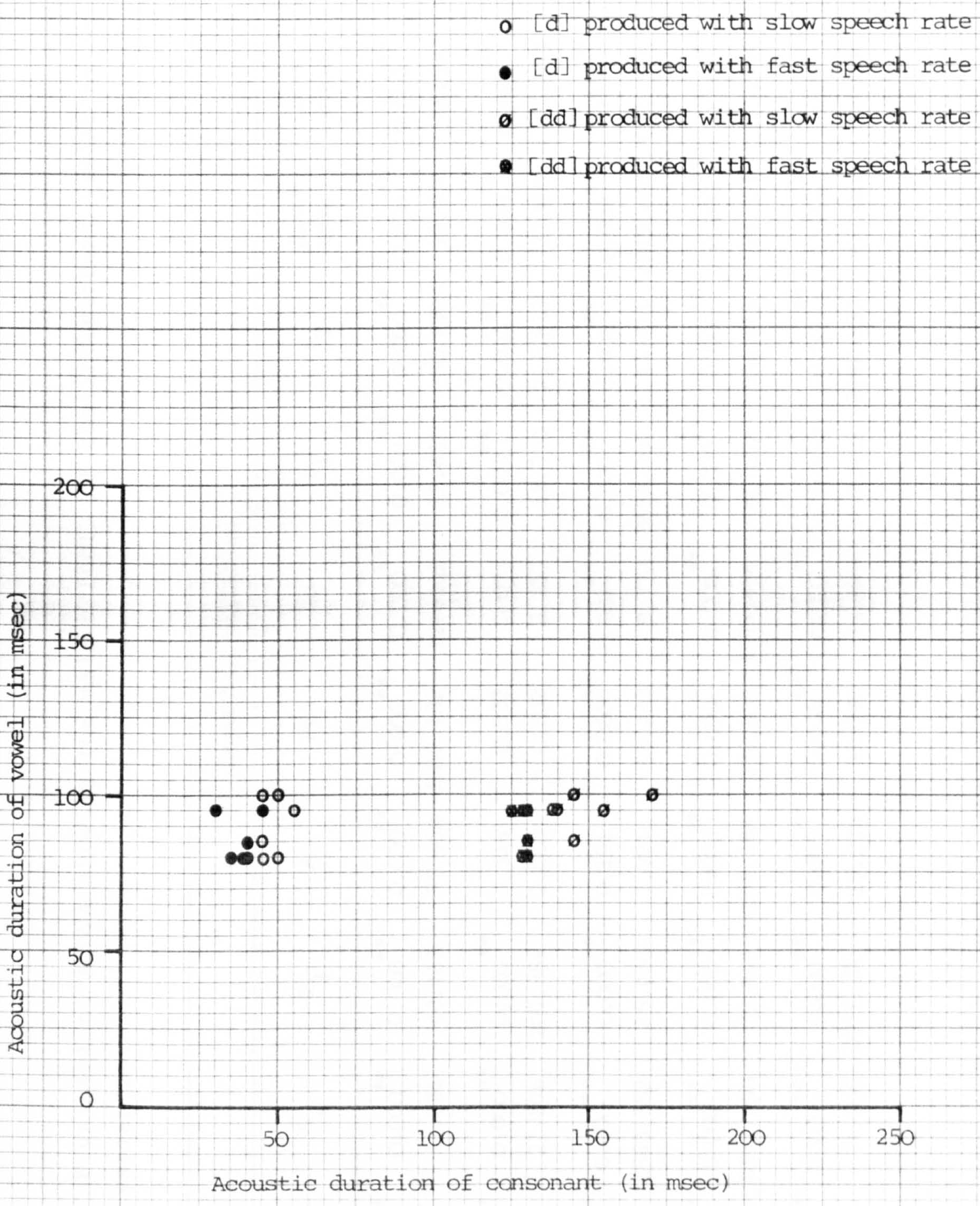


Fig. (6.16) Graph showing acoustic durations for single and geminate plosive consonants (Subject : AM)

Table (6.46)

Results (p and F values) of ANOVA showing the main effects of the independent variables on the dependent variables.

			Dependent Variables								
Independent Variables	Main Effects	Values	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2	Arvc-ratio	Acvc-ratio
	SPKR	p	0.001**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**
		F	10.843	21.265	43.163	77.450	1012.132	226.212	12.651	50.270	19.704	18.690
	RATE	p	0.000**	0.000**	0.000**	0.316	0.532	0.073	0.000**	0.000**	0.000**	0.000**
		F	30.366	21.876	207.859	1.010	0.392	3.267	18.770	211.913	73.871	65.934
	POSN	p	0.000**	0.000**	0.000**	0.191	0.000**	0.002**	0.000**	0.001**	0.000**	0.000**
		F	****	2563.352	16.430	1.721	58.433	9.770	8651.484	11.631	4819.660	4782.516
	TYPE	p	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**
		F	31.392	2563.400	32.944	1854.686	278.509	2273.434	60.072	504.665	24.325	169.344
	SINGEM	p	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.119	0.000**	0.000**	0.000**
		F	17.291	71.205	2667.120	17.231	52.377	12.899	2.462	2198.957	852.068	507.406

Table (6.47)

Results (p and F values) of ANOVA showing the 2-way interactions effects of the independent variables on the dependent variables.

			Dependent Variables									
	2-way Interactions	Values	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2	Arvc-ratio	Acvc-ratio
Independent Variables	SPKR	p	0.027*	0.002**	0.023*	0.525	0.000**	0.344	0.538	0.025*	0.125	0.525
	RATE	F	4.952	9.574	5.241	0.406	62.016	0.901	0.381	5.113	2.375	0.406
	SPKR	p	0.001**	0.000**	0.203	0.000**	0.000**	0.000**	0.000**	0.267	0.000**	0.000**
	POSN	F	10.833	21.261	1.632	14.427	17.484	24.646	12.661	1.241	20.769	18.935
	SPKR	p	0.000**	0.000**	0.106	0.000**	0.000**	0.000**	0.368	0.000**	0.207	0.001**
	TYPE	F	31.420	21.265	2.650	77.450	28.596	226.212	0.813	28.073	1.602	11.367
	SPKR	p	0.387	0.014*	0.441	0.731	0.108	0.932	0.476	0.864	0.348	0.179
	SINGEM	F	0.753	6.198	0.596	0.119	2.611	0.007	0.512	0.029	0.884	1.819
	RATE	p	0.000**	0.000**	0.505	0.652	0.975	0.618	0.000**	0.359	0.000**	0.000**
	POSN	F	30.348	21.872	0.446	0.205	0.001	0.250	18.759	0.848	75.726	67.926

Table (6.47) cont'd

			Dependent Variables									
	2-way Interactions	Values	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2	Arvc-ratio	Acvc-ratio
Independent Variables	RATE	p	0.103	0.000**	0.656	0.316	0.009**	0.073	0.201	0.015*	0.881	0.116
	TYPE	F	2.695	21.876	0.200	1.010	7.022	3.267	1.650	6.078	0.022	2.500
	RATE	p	0.084	0.000**	0.000**	0.742	0.416	0.828	0.964	0.000**	0.000**	0.000**
	SINGEM	F	3.022	16.068	88.801	0.109	0.664	0.048	0.002	84.310	27.690	26.990
	POSN	p	0.000**	0.000**	0.005**	0.191	0.000**	0.002**	0.000**	0.011*	0.000**	0.000**
	TYPE	F	31.374	2563.352	8.033	1.721	37.257	9.770	60.093	6.561	24.014	180.705
	POSN	p	0.000**	0.000**	0.871	0.242	0.001**	0.497	0.118	0.000**	0.000**	0.000**
	SINGEM	F	17.304	71.213	0.026	1.379	10.452	0.464	2.467	13.425	890.267	537.212
	TYPE	p	0.067	0.000**	0.222	0.000**	0.644	0.000**	0.420	0.000**	0.320	0.027*
	SINGEM	F	3.407	71.205	1.502	17.231	0.215	12.899	0.654	18.509	0.996	5.000

Table (6.48)

Results (p and F values) of ANOVA showing the 3-way interactions effects of the independent variables on the dependent variables.

Independent Variables			Dependent Variables							
	3-Way Interactions	Values	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2
	SPKR RATE	p	0.028^{*}	0.002^{**}	0.693	0.619	0.002^{**}	0.851	0.539	0.312
	POSN	F	4.945	9.571	0.157	0.248	9.628	0.035	0.379	1.027
	SPKR RATE	p	0.068	0.002^{**}	0.693	0.525	0.038^{*}	0.344	0.170	0.310
	TYPE	F	3.380	9.574	0.157	0.406	4.364	0.901	1.898	1.039
	SPKR RATE	p	0.083	0.003^{**}	0.001^{**}	0.049^{*}	0.020^{*}	0.070	0.671	0.000^{**}
	SINGEM	F	3.049	9.218	11.682	3.951	5.485	3.338	0.181	16.411
	SPKR POSN	p	0.000^{**}	0.000^{**}	0.609	0.000^{**}	0.000^{**}	0.000^{**}	0.639	0.263
	TYPE	F	31.402	21.261	0.263	14.427	13.831	24.646	0.811	1.261
	SPKR POSN	p	0.386	0.014^{*}	0.383	0.641	0.701	0.653	0.475	0.019^{*}
	SINGEM	F	0.756	6.201	0.767	0.219	0.148	0.203	0.514	5.597

Table (6.48) cont'd

			Dependent Variables							
	3-Way Interactions	Values	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2
Independent Variables	SPKR TYPE SINGEM	p	0.006^{**}	0.014^{*}	0.258	0.731	0.128	0.932	0.419	0.071
		F	7.874	6.198	1.288	0.119	2.341	0.007	0.658	3.293
	RATE POSN TYPE	p	0.103	0.000^{**}	0.505	0.652	0.268	0.618	0.200	0.568
		F	2.689	21.872	0.446	0.205	1.236	0.250	1.653	0.328
	RATE POSN SINGEM	p	0.084	0.000^{**}	0.861	0.324	0.036^{*}	0.569	0.963	0.473
		F	3.028	16.072	0.031	0.980	4.458	0.326	0.002	0.516
	RATE TYPE SINGEM	p	0.084	0.000^{**}	0.120	0.742	0.393	0.828	0.740	0.205
		F	3.014	16.068	2.443	0.109	0.735	0.048	0.111	1.617
	POSN TYPE SINGEM	p	0.067	0.000^{**}	0.005^{**}	0.242	0.154	0.497	0.421	0.945
		F	3.413	71.213	8.255	1.379	2.047	0.464	0.652	0.005

Table (6.49)

Results (p and F values) of ANOVA showing the 4-way interactions effects of the independent variables on the dependent variables.

			Dependent Variables							
	4-Way Interactions	Values	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2
Independent Variables	SPKR RATE	p	0.068	0.002^{**}	0.032^{*}	0.619	0.661	0.851	0.170	0.261
	POSN TYPE	F	3.375	9.571	4.664	0.248	0.194	0.035	1.902	1.273
	SPKR RATE	p	0.082	0.003^{**}	0.861	0.062	0.423	0.053	0.672	0.415
	POSN SINGEM	F	3.054	9.221	0.031	3.532	0.645	3.798	0.180	0.669
	SPKR RATE	p	0.004^{**}	0.003^{**}	0.618	0.049^{*}	0.113	0.070	0.417	0.567
	TYPE SINGEM	F	8.415	9.218	0.249	3.951	2.545	3.338	0.662	0.328
	SPKR POSN	p	0.006^{**}	0.014^{*}	0.397	0.641	0.259	0.653	0.418	0.413
	TYPE SINGEM	F	7.883	6.201	0.722	0.219	1.284	0.203	0.660	0.673
	RATE POSN	p	0.084	0.000^{**}	0.390	0.324	0.380	0.569	0.739	0.247
	TYPE SINGEM	F	3.019	16.072	0.744	0.980	0.775	0.326	0.111	1.351

Table (6.50)

Results (p and F values) of ANOVA showing the 5-way interactions effects of the independent variables on the dependent variables.

			Dependent Variables							
	5-way Interactions	Values	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2
Independent variables	SPKR RATE POSN TYPE SINGEM	p	0.004**	0.003**	0.085	0.062	0.287	0.053	0.416	0.597
		F	8.424	9.221	2.999	3.532	1.141	3.798	0.664	0.281

Table (6.51)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 1, Type = 1, Single vs. Geminate = 1) are kept constant.

	Variables				
Groups	T_2	Ac	Po	Uo	t_2
Rate	0.113	0.910	0.001**	0.532	0.003**

Table (6.52)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 1, Type = 1, Single vs. Geminate = 1) are kept constant.

	Variables				
Groups	T_2	Ac	Po	Uo	t_2
Rate	0.295	0.666	0.353	0.691	0.106

Table (6.53)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 1, Type = 1, Single vs. Geminate = 2) are kept constant.

	Variables				
Groups	T_2	Ac	Po	Uo	t_2
Rate	0.001**	0.141	0.370	0.052	0.001**

Table (6.54)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 1, Type = 1, Single vs. Geminate = 2) are kept constant.

	Variables				
Groups	T_2	Ac	Po	Uo	t_2
Rate	0.000**	0.372	0.263	0.421	0.002**

Table (6.55)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 2, Type = 1, Single vs. Geminate = 1) are kept constant.

	Variables									
Groups	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2	Arvc-ratio	Acvc-ratio
Rate	0.168	0.297	0.290	0.250	0.002**	0.755	0.405	0.018*	0.787	0.203

Table (6.56)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 2, Type = 1, Single vs. Geminate = 1) are kept constant.

	Variables									
Groups	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2	Arvc-ratio	Acvc-ratio
Rate	0.428	0.756	0.372	0.081	0.071	0.170	0.496	0.160	0.926	0.527

Table (6.57)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 2, Type = 1, Single vs. Geminate = 2) are kept constant.

	Variables									
Groups	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2	Arvc-ratio	Acvc-ratio
Rate	0.000^{**}	0.000^{**}	0.002^{**}	0.349	0.019^{*}	0.043^{*}	0.288	0.000^{**}	0.955	0.003^{**}

Table (6.58)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 2, Type = 1, Single vs. Geminate = 2) are kept constant.

	Variables									
Groups	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2	Arvc-ratio	Acvc-ratio
Rates	0.481	0.424	0.015^{*}	0.241	0.119	0.163	0.022^{*}	0.021^{*}	0.028^{*}	0.256

Table (6.59)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 1, Type = 2, Single vs. Geminate = 1) are kept constant.

	Variables		
Groups	T_2	Po	t_2
Rate	0.002**	0.156	0.599

Table (6.60)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 1, Type = 2, Single vs. Geminate = 1) are kept constant.

	Variables		
Groups	T_2	Po	t_2
Rate	0.014*	0.063	0.000**

Table (6.61)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 1, Type = 2, Single vs. Geminate = 2) are kept constant.

	Variables		
Groups	T_2	Po	t_2
Rate	0.001^{**}	0.874	0.000^{**}

Table (6.62)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 1, Type = 2, Single vs. Geminate = 2) are kept constant.

	Variables		
Groups	T_2	Po	t_2
Rate	0.016^{*}	0.718	0.021^{*}

Table (6.63)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 2, Type = 2, Single vs. Geminate = 1) are kept constant.

	Variables						
Groups	T_1	T_2	Po	t_1	t_2	Arvc-ratio	Acvc-ratio
Rate	0.020^{*}	0.013^{*}	0.023^{*}	0.000^{**}	0.551	0.426	0.063

Table (6.64)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 2, Type = 2, Single vs. Geminate = 1) are kept constant.

	Variables						
Groups	T_1	T_2	Po	t_1	t_2	Arvc-ratio	Acvc-ratio
Rate	0.271	0.001^{**}	0.001^{**}	0.398	0.007^{**}	0.005^{**}	0.069

Table (6.65)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 2, Type = 2, Single vs. Geminate = 2) are kept constant.

	Variables						
Groups	T_1	T_2	Po	t_1	t_2	Arvc-ratio	Acvc-ratio
Rate	0.040^*	0.000^{**}	0.590	0.038^*	0.000^{**}	0.000^{**}	0.002^{**}

Table (6.66)

T-test results (p values) showing the effect of 'Rate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 2, Type = 2, Single vs. Geminate = 2) are kept constant.

	Variables						
Groups	T_1	T_2	Po	t_1	t_2	Arvc-ratio	Acvc-ratio
Rate	0.075	0.004^{**}	0.385	0.071	0.023^*	0.080	0.540

Table (6.67)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variable when all other independent variables (i.e. Speaker = 1, Position = 1, Type = 1, Rate = 1) are kept constant.

	Variables				
Groups	T_2	Ac	Po	Uo	t_2
Single vs. Geminate	0.000**	0.315	0.028*	0.430	0.000**

Table (6.68)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 1, Type = 1, Rate = 1) are kept constant.

	Variables				
Groups	T_2	Ac	Po	Uo	t_2
Single vs. Geminate	0.000**	0.019*	0.071	0.020*	0.000**

Table (6.69)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 1, Type = 1, Rate = 2) are kept constant.

	Variables				
Groups	T_2	Ac	Po	Uo	t_2
Single vs. Geminate	0.000^{**}	0.006^{**}	0.296	0.027^{*}	0.000^{**}

Table (6.70)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 1, Type = 1, Rate = 2) are kept constant.

	Variables				
Groups	T_2	Ac	Po	Uo	t_2
Single vs. Geminate	0.000^{**}	0.032^{*}	0.007^{**}	0.046^{*}	0.003^{**}

Table (6.71)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 2, Type = 1, Rate = 1) are kept constant.

	Variables									
Groups	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2	Arvc-ratio	Acvc-ratio
Single vs. Geminate	0.001^{**}	0.000^{**}	0.000^{**}	0.052	0.220	0.037^{*}	0.288	0.000^{**}	0.001^{**}	0.000^{**}

Table (6.72)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 2, Type = 1, Rate = 1) are kept constant.

	Variables									
Groups	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2	Arvc-ratio	Acvc-ratio
Single vs. Geminate	0.793	0.059	0.000^{**}	0.444	0.913	0.468	0.878	0.000^{**}	0.000^{**}	0.002^{**}

Table (6.73)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 2, Type = 1, Rate = 2) are kept constant.

	Variables									
Groups	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2	Arvc-ratio	Acvc-ratio
Single vs. Geminate	0.580	0.061	0.000^{**}	0.796	0.262	0.858	0.405	0.000^{**}	0.000^{**}	0.000^{**}

Table (6.74)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 2, Type = 1, Rate = 2) are kept constant.

	Variables									
Groups	T_1	T_4	T_2	Ac	Po	Uo	t_1	t_2	Arvc-ratio	Acvc-ratio
Single vs. Geminate	0.399	0.020^{*}	0.000^{**}	0.045^{*}	0.000^{**}	0.066	0.791	0.000^{**}	0.000^{**}	0.000^{**}

Table (6.75)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 1, Type = 2, Rate = 1) are kept constant.

	Variables		
Groups	T_2	Po	t_2
Single vs. Geminate	0.000**	0.000**	0.000**

Table (6.76)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 1, Type = 2, Rate = 1) are kept constant.

	Variables		
Groups	T_2	Po	t_2
Single vs. Geminate	0.000**	0.237	0.000**

Table (6.77)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 1, Type = 2, Rate = 2) are kept constant.

	Variables		
Groups	T_2	Po	t_2
Single vs. Geminate	0.000**	0.002**	0.000**

Table (6.78)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 1, Type = 2, Rate = 2) are kept constant.

	Variables		
Groups	T_2	Po	t_2
Single vs. Geminate	0.000**	0.007**	0.000**

Table (6.79)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 2, Type = 2, Rate = 1) are kept constant.

	Variables						
Groups	T_1	T_2	Po	t_1	t_2	Arvc-ratio	Acvc-ratio
Single vs. Geminate	0.599	0.000^{**}	0.092	0.726	0.000^{**}	0.000^{**}	0.000^{**}

Table (6.80)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 2, Type = 2, Rate = 1) are kept constant.

	Variables						
Groups	T_1	T_2	Po	t_1	t_2	Arvc-ratio	Acvc-ratio
Single vs. Geminate	0.064	0.000^{**}	0.130	0.242	0.000^{**}	0.000^{**}	0.000^{**}

Table (6.81)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 1, Position = 2, Type = 2, Rate = 2) are kept constant.

	Variables						
Groups	T_1	T_2	Po	t_1	t_2	Arvc-ratio	Acvc-ratio
Single vs. Geminate	0.414	0.000**	0.614	0.111	0.000**	0.000**	0.000**

Table (6.82)

T-test results (p values) showing the effect of 'Single vs. Geminate' on the dependent variables when all other independent variables (i.e. Speaker = 2, Position = 2, Type = 2, Rate = 2) are kept constant.

	Variables						
Groups	T_1	T_2	Po	t_1	t_2	Arvc-ratio	Acvc-ratio
Single vs. Geminate	0.080	0.000**	0.014*	0.570	0.000**	0.000**	0.000**

Table (6.83)

T-test results (p values) showing the degree of interaction between a single fricative spoken with slow speaking rate and its geminate cognate spoken with fast speaking rate in word-initial position.

Subject	T_2	t_2
1	0.000^{**}	0.001^{**}
2	0.000^{**}	0.009^{**}

Table (6.84)

T-test results (p values) showing the degree of interaction between a single fricative spoken with slow speaking rate and its geminate cognate spoken with fast speaking rate in word-medial position.

Subject	T_2	t_2
1	0.000^{**}	0.000^{**}
2	0.000^{**}	0.000^{**}

Table (6.85)

T-test results (p values) showing the degree of interaction between a single plosive spoken with slow speaking rate and its geminate cognate spoken with fast speaking rate in word-initial position.

Subject	T_2	t_2
1	0.000^{**}	0.000^{**}
2	0.001^{**}	0.001^{**}

Table (6.86)

T-test results (p values) showing the degree of interaction between a single plosive spoken with slow speaking rate and its geminate cognate spoken with fast speaking rate in word-medial position.

Subject	T_2	t_2
1	0.000^{**}	0.000^{**}
2	0.000^{**}	0.000^{**}

Table (6.87)

T-test results (p values) showing the degree of interaction between a single fricative spoken with slow speaking rate and its geminate cognate spoken with fast speaking rate in word-initial and word-medial positions.

Subject	T_2	t_2
1	0.000^{**}	0.000^{**}
2	0.000^{**}	0.000^{**}

Table (6.88)

T-test results (p values) showing the degree of interaction between a single plosive spoken with slow speaking rate and its geminate cognate spoken with fast speaking rate in word-initial and word-medial positions.

Subject	T_2	t_2
1	0.000^{**}	0.000^{**}
2	0.000^{**}	0.000^{**}

Table (6.89)

T-test results (p values) showing the degree of interaction between single fricative and plosive consonants spoken with slow speaking rate and their geminate cognates spoken with fast speaking rate in word-initial and word-medial positions.

Subject	T_2	t_2
1	0.000^{**}	0.000^{**}
2	0.000^{**}	0.000^{**}

CHAPTER SEVEN

CONCLUSIONS AND SUGGESTIONS FOR FURTHER INVESTIGATIONS

7.1 Conclusions

The foregoing chapters of this study are an attempt to elucidate certain phonetic and phonological aspects of the geminate/non-geminate contrast in I.C. Arabic. The final conclusions that could be drawn from the data indicate that duration plays a vital role in distinguishing single consonants from geminate consonants. An examination of the durational contrasts between single and geminate consonants has shown that geminate consonants are always considerably longer than their non-geminate partners whether they are produced in isolated words or in context. Our results strongly suggest that consonant duration is regarded as a major factor in the linguistic functioning of gemination.

Differences in speech rate have shown some very interesting effects on the distinction between single and geminate consonants in I.C. Arabic. For instance, a geminate consonant is not only considerably longer than its non-geminate cognate when the two segments are articulated either in words spoken with slow speech rate or in words spoken with fast speech rate, but also a geminate consonant produced in words spoken with fast speaking rate is significantly longer than a single counterpart produced in words spoken with slow speaking rate. This quite obviously indicates that temporal overlapping between a geminate consonant and its single cognate never occurs under any circumstances. Further, geminate consonants used in words uttered with slow speaking rate are significantly longer than those used in words uttered with fast speaking rate. Single consonants have also been shown to be longer when produced with slow speech rate than when produced with fast speech rate. Long segments produced with slow speech rate suffer major reductions in their durations when they are produced with fast rate of speech. On the other hand, segments of short duration are less affected by differences in speech rate. This suggests that the longer the sound segment is the greater reduction in its duration it suffers when it is produced at rapid rates of speech. Since geminate consonants are longer than single consonants, they are then liable to be more severely shortened at fast speech rates than their single partners.

In contrast with previous studies (see section 5.13.1), our present study presents evidence that not only vowels are shortened when speech rate is increased but also consonants. Our current data bear testimony to the fact that the consonantal portion within an utterance is more affected by an increase in speech rate than the vocalic portion. The vowels lose only a relatively small amount of their original durations when produced with fast rate of speech; whereas the consonants shorten more than the vowels when speech rate increases. Vowels occurring in stressed positions maintain their original length when they precede or follow an intervocalic geminate consonant used in words spoken in isolation or in contextual utterance.

In distinguishing a geminate consonant from a single one, the duration of the preceding vowel is a negligible factor. Vowels preceding single consonants are not significantly different from those preceding geminate consonants except, in articulatory timing, before fricatives. The insignificant acoustic durational differences between a vowel preceding a geminate consonant and that preceding a single consonant could be ascribed to the fact that the friction noise of a fricative or the closure time of a plosive for both the geminate and the non-geminate virtually starts at the same time but it lasts predominantly longer with geminate than with non-geminate consonants. In addition, our data do not provide evidence in support of the idea of temporal compensation. The measurements do not show any evidence of existing compensation between the lengthening of an intervocalic consonant and a preceding vowel; and they show that durational differences between vowels preceding geminate consonants and those preceding non-geminate consonants are negligible. Neither on the acoustic nor on the articulatory levels is there in our data any indication of the existence of temporal compensatory adjustment between the preceding vowel and the following geminate/non-geminate consonant.

From the aerodynamic point of view, geminate consonants are generally accompanied by higher intraoral pressure than their non-geminate partners no matter whether they are occurring word-initially or word-medially, produced by one and the same speaker or by different speakers, used in words said with slow speech rate or said with fast speech rate. These findings have been shown to give further support to previous findings reported by Hixon (1966). Also, oral airflow is greater for single fricative consonants than for their geminate counterparts, and minimum cross-sectional area of constriction is significantly smaller for a geminate fricative than for a single cognate. The considerable differences in oral airflow and oral pressure between a single [s] and a geminate [ss] are associated with the degree of constriction in the supraglottal vocal tract. But there could be other factors operating too, at the glottis, or in the subglottal articulators. There is a closer approximation between the articulating organs, namely the tip or blade of the tongue and the alveolar ridge, during the articulation of [ss] than during the articulation of [s]. In consequence, this results in greater intraoral pressure and lower airflow rate for the geminate and lesser intraoral pressure and higher airflow rate for the non-geminate. At the same time, evidence has been provided indicating that there is higher intraoral pressure during the articulation of single and geminate voiceless fricatives than during the articulation of single and geminate voiced plosives. Such a finding signifies that the closure of the glottis, a necessary requisite to produce voicing, increases the resistance within the vocal tract and, accordingly, reduces the flow rate.

Articulatorily speaking, evidence has been presented showing that the tongue touches the palate more firmly and over a greater contact area during the production of geminate consonants than during the production of non-geminate ones. The firmer contact between the articulating organs is inferred from a clearer and more extensive removal of the marking medium, as shown on the palatograms, for the geminates than for the non-geminates. The greater mechanical pressure and the larger area of contact are probably associated with the longer duration of contact for the geminates. The clearer and greater wipe-off associated with the geminate consonants is presumably a natural outcome of a greater pressure exerted by the tongue against the palate during the articulation of the geminates than during the articulation of the non-geminates.

Moreover, it has been found that intraoral air pressures and airflow rates are consistently higher for voiceless consonants than for their voiced cognates regardless of differences in speech rate as well as differences in speakers. Intraoral pressure, oral airflow and minimum cross-sectional area of supraglottal constriction do not appear to change significantly with changes in rate of speech, suggesting that geminate and non-geminate consonants produced with fast speech rate show insignificant differences in oral pressure, oral airflow and minimum cross-sectional area of constriction from geminate and non-geminate consonants produced with slow speech rate. We, therefore, assume that during increased speed of geminate versus non-geminate articulations the manner in which the investigated consonants are produced does not change significantly. Also, because oral pressure and oral airflow do not change with changing speech rate, we find it reasonable to assume that the shape of the supraglottal constriction may not have changed at different speaking rates.

From the perceptual point of view, our speech synthesis experiments have shown that listeners could easily distinguish geminate consonants from their single counterparts when the duration of the preceding vowel remained unchanged. The perceptual tests confirmed that consonant duration is a major cue for the perception of gemination; the duration of the preceding vowel does not appear to contribute to the perceptual realization of the following long consonant as a geminate consonant. The durational differences in geminate/non-geminate consonants are far beyond the DL values and are, therefore, perceptually significant for the listeners judgment of the geminates as long consonants; hence they are phonologically distinctive. On the other hand, the insignificant durational differences of vowels preceding geminate/non-geminate consonants hover around the DL values. They are then neither phonetically nor phonologically significant.

Perhaps the most interesting conclusion drawn from this study is that gemination in Arabic does not involve a rearticulation of the same consonant. Phonetically, there is no evidence that the production of word-initial or word-medial geminate consonant involves double articulation. It is produced as one long indivisible consonant. Besides, our data have suggested that duration is a likely, but probably not the only, candidate for an invariant acoustic property to distinguish between single and geminate consonants in I.C. Arabic.

7.2 Suggestions for Further Investigations

Our present study is the most comprehensive experimental investigation of consonant gemination in Arabic to date. There still remain a number of points that need to be further investigated so as to clarify the overall picture and shed more light on the single/geminate contrast in Arabic.

It would be interesting if further research examined the distinction between voiced and voiceless geminates to find out what really differentiates between a voiced geminate and a voiceless partner, especially in so far as the articulatory and the aerodynamic processes are concerned. It would also be interesting if a more sophisticated experiment on synthetic speech is conducted, provided that it should aim at pursuing whether the single/geminate contrast reflects a genuine case of the phenomenon of categorical perception. In this case one would not only need to conduct parallel experiments using the same duration continuum to test for peaks in the discrimination function, but he would also need to do further synthetic experiments to illustrate non-categorical perception.

Geminate consonants can be emphatic (i.e. mufaxxama) as well as non-emphatic. To assess the distinction between an emphatic geminate and a non-emphatic counterpart, an articulatory investigation in which cinefluorography is employed is required. The obtained X-ray motion pictures should show the articulatory movements of the larynx as well as the other speech organs above the larynx. Furthermore, an electromyographic investigation is of necessity to investigate whether an extra muscle activity is involved in the production of geminate consonants.

It would be of interest to register the timing of glottal opening and closing gestures in connection with the opening and closing gestures of the supraglottal constriction while producing geminate consonants. To achieve this goal, one would need to use either photo-electric glottography to measure the relative amount of light transmitted through the glottis, or electro-laryngography to measure the relative conductance or impedance across the vocal folds. The latter technique measures the duration of vocal fold closure in each vibratory cycle, but it tells nothing about the width or shape of vocal fold opening; whereas the former technique allows us to obtain information on the area of glottal opening as an indirect measure of vocal fold adjustment. It gives no information on the shape of the glottis but only on the extent of the opening. (Borden and Harris, 1980). These two techniques would be of great value in further investigation of the Arabic pharyngeal geminates.

We still believe that a more sophisticated experiment using electro-palatography is of ultimate necessity to examine tongue-palate contact during the articulation of geminate/non-geminate consonants. This will greatly help to confirm the findings we have presented in our direct palatography experiment.

Finally, it must be recognized that, although there are great practical difficulties in obtaining subjects for physiological experiments, one would ideally wish to work with a larger group of subjects.

APPENDICES

Appendix (1)

SPS256

A Program for Speech Synthesis with a PET Computer and a Computalker Speech Synthesiser.

INSTRUCTION MANUAL

Written by Janak Gandhi and Peter Roach

January 1980

Department of Linguistics and Phonetics

University of Leeds

1. Introduction

This manual explains how to use the SPS256 program to synthesise speech. The general principles are explained in this section, and the detailed operating instructions are in the following sections. The program has been designed to let you gain experience in manipulating acoustic phonetic parameters as simply as possible; however, you must be clear that speech synthesis is a slow and laborious job that requires patience, and that the computer will only work properly if you use it according to the instructions.

The Computalker synthesiser has nine parameters, and we begin by explaining what these are. Firstly, there are four amplitude parameters relating to the four basic kinds of sound the synthesiser can produce: AV for vowels and vowel-like sounds, AH for aspiration, AF for fricative noise and AN for a nasal sound. Any value for any of these must be in the range 0 to 225, where 255 is the maximum amplitude. The figures do not correspond to any physical scale. If we want a vowel, the value of AV must be high enough to produce audible vowel-type sound and the other three amplitudes must be at zero. If on the other hand we want a voiceless fricative other than [h] we must have AF above zero and the other three amplitudes at zero. We can, of course, combine these sounds by having more than one amplitude parameter above zero, but we will not discuss this here in order to keep things simple.

AV and AN are parameters that involve voicing, while AF and AH involve noise. Any voiced sound must have a fundamental frequency, so we can now introduce another parameter, FO. The values for this parameter are in Hz, with a range from 74 to 464 Hz.

The remaining four parameters relate to spectral characteristics. For vowels and vowel-like sounds it is necessary to have values for the first three formants. The ranges are as follows (in Hz):

Parameter	Minimum	Maximum
F1	175	1452
F2	527	4356
F3	1704	5508

Finally, there is a fricative formant FF, with a frequency range of 1705 to 14160 Hz.

To produce reasonably clear speech, the synthesiser must be fed with values for each parameter at least 100 times per second. To provide these 900 numbers per second manually would be impossibly lengthy even if you could do it fast enough; it is the computer's job to produce these very large quantities of numbers, and it is the user's job to give the computer the information it needs in order to generate the correct numbers. Let us now look at some simple examples to see how this is done.

If we wish to produce a steady vowel as in Fig. 1, clearly all we need to do is give the appropriate values to the computer and tell it how long to keep sending these values to the synthesiser every hundredth of a second (or whichever other "frame rate" is chosen). Thus we could add to the nine parameter values a tenth specifying the duration, so that the data for this vowel might look this this:

Table 1

AV	200	(an arbitrary choice)
FO	100	
F1	500	
F2	1500	
F3	2500	
AH	0	
AF	0	
FF	5000	(frequency parameters must always have a value in their appropriate range)
AN	0	
PD	200	(duration in msec)

Now let us consider how we would produce a vowel glide as in Fig. 2. Obviously we want to let the computer do the calculation of the values intermediate between point 1 and point 2. We could do this by giving the computer two sets of data, one for the starting values (plus the duration of the vowel glide) and the other for the finishing values. Since these finishing values are only to be reached at the end, we make the duration (PD) for point 2 zero. So we give the computer the values in Table 2 and let the computer work out the numbers to be fed to the synthesiser.

Table 2

Point 1		Point 2	
AV	200	AV	200
FO	100	FO	100
F1	800	F1	250
F2	1500	F2	2000
F3	3000	F3	3000
AH	0	AH	0
AF	0	AF	0
FF	5000	FF	5000
AN	0	AN	0
PD	200	PD	0

We now have a problem. If the computer is programmed to work out a straight line from point to point as in the above example, how could we produce a sudden jump as in Fig. 3? The program has been designed to make this possible. Since minus values of parameters are impossible, we give a minus sign to any parameter which is to stay as it is until the next point. It will then stay steady, and if the value at the next point is different it will do a sudden jump to that value. So the data to produce an output like Fig. 3 will be as in Table 3.

Table 3

Point 1		Point 2	
AV	200	AV	200
F0	100	F0	100
F1	- 800	F1	250
F2	-1500	F2	2000
F3	3000	F3	3000
AH	0	AH	0
AF	0	AF	0
FF	5000	FF	5000
AN	0	AN	0
PD	200	PD	0

To end the sounds clearly, the computer produces silence after transmitting the last values. Thus in the above example, once the last values are transmitted the vowel amplitude (AV) will suddenly drop from 200 to 0. If we want the sound to decay away to silence gradually, we need to add Point 3 at the end, in which AV has the value 0. When we give a PD value of zero for a point, that point will in fact have a duration of a single "frame" - for example, if the user selects a rate of 100 frames per second (intervals of 10 msec), a PD value of zero will result in a duration of 10 msec.

Point 1		Point 2		Point 3	
AV	200	AV	200	AV	0
FO	100	FO	100	FO	100
F1	- 800	F1	250	F1	250
F2	-1500	F2	2000	F2	2000
F3	2500	F3	2500	F3	2500
AH	0	AH	0	AH	0
AF	0	AF	0	AF	0
FF	5000	FF	5000	FF	5000
AN	0	AN	0	AN	0
PD	200	PD	100	PD	0

A lot of practical experience is necessary to produce convincing sounding synthetic speech, but the basic principle of this system should now be clear: it is that the user decides the points in time at which important changes in acoustic structure take place, and whether there is a gradual or an abrupt change from the values of one point to the values of the following point. How many points are used depends to some extent on how precise the specification of the acoustic structure of the speech is to be. On average perhaps three or four points per phoneme are necessary.

We will conlcude the introduction with a brief summary of how we do speech synthesis.

(i) The user decides what speech is to be synthesised. Almost always this will involve working from a spectrogram that has been segmented and measured.

(ii) The data for each point is entered on the coding sheet.

(iii) The computer, synthesiser and related equipment is switched on, the program loaded and set in operation.

(iv) Point data is typed into the computer's memory, and the computer creates the data frames.

(v) The user can examine the data in its final forms on graphs on the computer screen or printer, and can listen to the result.

(vi) If the Ferrograph Logic 7 tape recorder is connected, the speech can be automatically recorded.

(vii) The point data can be stored in digital form in cassette for later re-use.

2. Operating Instructions

As described in the previous section, to synthesise speech the basic role of the computer is to read parameter values at discrete points, calculate intermediate values and send them to the synthesiser at a specified rate. The program SPS256, besides doing the above, also provides facilities for editing, displaying and saving data.

The program has been designed to simplify the user's task as much as possible. Whenever interaction with the user is needed, the program causes messages to be displayed and most of them are self-explanatory.

Basically, the program is a monitor; it accepts the user's commands from the keyboard, carries out the required job and waits for the next command. All commands that it can accept are abbreviated to a single letter and are explained later. Also the parameter names have been abbreviated to two characters.

Depending on the status of the program, certain commands or even the data values may not be accepted, in which case a message explaining the reason for rejection is displayed and the program waits either for the next command or a new data value. Whenever a '*' is displayed on the screen it indicates that it is ready and waiting for a command. When a '?' is displayed, it means that the user is being asked to input a data value or a parameter name as specified by whatever precedes '?' on the screen.

We now look at how we make the program ready for the first time. To do this, the program must be loaded in the computer's memory and activated. The following steps are required:

i) switch on the computer, the synthesiser and the Ferrograph

ii) load the tape no. I on tape unit with SPS256 side up

iii) type: LOAD "SPS256" and press RETURN key

iv) press PLAY of tape unit when the message

PRESS PLAY ON TAPE 1 appears on screen

v) READY will be displayed when the program has been loaded into the memory. Rewind tape and take it out.

vi) type: RUN and press RETURN. The screen will show

```
* THIS IS THE SPEECH SYNTHESIS PROGRAM ENTER COMMAND
*
```

and the program is ready to accept your command.

The command should be just its abbreviation, it should not be followed by RETURN. On the other hand, each data value or name should be terminated by RETURN.

If at any stage, following a command, RETURN is pressed by mistake and the program halts, type in

GOTO 180

The execution will restart and the command can be entered again.

Now follows the description of various commands, their functions and requirements.

I INPUT of data to computer.

Once the program is activated this has to be the first command to feed data to the program. Also when the user has finished with one set of data, it makes it possible to read in fresh data for new speech production, without reloading or restarting the program.

The program can accept data either from keyboard or from digital tape. When the command I is typed in, the computer will ask

WHICH SOURCE?

Type in T or K depending on whether the data is on tape or has to be entered from the keyboard. If tape is being used, the user must know the datafile name assigned at the file creation line, and enter it when asked for.

When fresh data is being entered, the program requires to know the number of points and frame interval before taking in data so it will display

NUMBER OF POINTS? and

FRAME INTERVAL?

successively.

There should be at least two data points and the maximum number can be 256. If the number does not lie in this range, an error message will be displayed and a new number will be asked for.

Frame interval is the duration between two consecutive frames. Generally it is 10 msec. Its value should always be greater than zero and should not exceed 45.

Now we turn to the actual data - the values of various parameters for each point, the point number is displayed, and then follow the parameter names one at a time followed by '?'. The values of the parameter name displayed is typed in, the next name will be displayed. It continues until all the values for each point have been read in. At the end a "*" will be displayed to indicate the end of input and the program waits for next command.

The value of the tenth parameter PD, the point duration, should be multiple of the frame interval, otherwise it will be rejected, and two nearest acceptable values will be displayed with an error message. The user should choose and enter a new value.

D DISPLAY of points data on screen.

The entered data can be inspected on the screen, before proceeding further. But because of the limited size of the screen, only five points can be displayed at a time. Thus points are divided in segments or groups of five points each. The last segment may have five or less points. When the command D is accepted, the screen will display

WHICH SEGMENT?

It will also be displayed after a segment whose number was typed in response is displayed. A segment number should be less than (INT (No. of pts./5) + 1) (INT (n) means integer part of number n) Segment number 0 will terminate the command.

L LISTS data for all the points on printer.

Before entering the command, switch on the printer if it is off. (Note: If command G has been executed before, switch off and then switch on the printer).

M MODIFY one parameter value for one point at a time.

The computer has to know what parameter, and which point. It will ask the user

WHICH POINT?

and WHICH PARAMETER?

successively. Then it will show the old value in the memory and asks for new value.

OLD VALUE -- NEW VALUE?

If PD is being modified, the new value should be a multiple of the frame interval.

O OVERALL modification.

This is a powerful command and saves the user from entering large quantities of data again. It can perform three functions depending upon the user's response to the question.

WHICH PARAMETER?

immediately following the command O.

The three possible responses and their functions are:

i) FI: Will change the value of frame interval. It will display the old value and request a new value.

The calculation of intermediate frame values depends on the size of FI. If FI is changed after the intermediate values have been calculated, it will be of no significance. But the program takes care to check that all PD's are multiples of new FI and also rejects the frame values if they have been created. If it finds a PD that is not a multiple of FI, its point number and two nearest feasible values will be displayed. So if the user is modifying FI, he must take care to observe messages and modify the invalid PD's using M command.

ii) NO: Modifies total number of points in memory. It can only append new points but cannot insert any point.

The computer will ask for -

NO. OF PT's TO BE ADDED?

Then checks that the total number does not exceed 256; if this happens, the number will be rejected and a new number will be asked for again. Having accepted the number of points to be appended, it will proceed to input parameter values as during the command I.

This command is useful in cases where the user has to enter data for a large number of points, then instead of entering all the data at one time it can be entered partially. Care must be taken that each time the number of points specified should be the number to be entered and not the total number.

iii) Any parameter name to modify all values of that parameter by a specified percentage. As soon as the parameter name is accepted the next question on screen will be

WHAT % AGE?

Enter the % age value by which the parameter has to be shifted throughout. Again if PD is being modified, the individual PD values appearing in error messages should be corrected.

S SAVE data on digital tape.

This command will save data on tape in the same order as is required by the program. Thus it can be used as input later.

Before saving the data the computer will ask you to give file name, and the data file will be assigned that name. When data has to be read as input, it can be accessed only by this filename. So the user must take care to keep a record of the filename and cassette no. used to save data.

When the computer prints "PRESS PLAY AND RECORD" please note that you must press RECORD before PLAY.

C CALCULATES the intermediate frame values.
This command will not be accepted if there exists some illegal values of PD and the message

MODIFY PD

will be displayed. User must check and modify PD's for all points, because partial modification will mask the error detection and the resulting speech not be accurate.

While it is busy in calculation the screen will show

GENERATING DATA FRAMES

until a '*' appears.
If a parameter value is out of its range or a last point parameter value should have a minus sign while its duration is more than FI, the program will be terminated by the error message.

ILLEGAL QUANTITY ERROR IN 1340

Then to restart the program type in

GOTO 180

and check all the values carefully and modify the error values.

The following commands are invalid until the command C has been executed, because they all need the intermediate values. If C has not been executed the error message

INVALID COMMAND, FRAMES NOT CREATED

will be displayed and control will be returned to monitor waiting for a command.

P PLAY will pass the data to the synthesiser, enabling the user to hear the speech. A period of silence is introduced between the moment of typing P and the beginning of the speech in order to allow the user to get ready to listen. The duration of the silence can be chosen and entered by the user when the screen displays

ENTER INITIAL SILENCE DURATION IN SECS

The user may wish to listen to the speech many times. This can be done by repeatedly typing P, but the user does not need to keep entering the silence duration. Once accepted, it is saved by the program until the user terminates the P command. The P command is terminated when the user types X (i.e. exit to monitor). Once the command P is started any command other than P will be invalid and will be rejected until X is typed.

R RECORDS the synthesised speech on the Ferrograph Logic 7. The program switches the tape recorder on and off before and after recording. To separate this sound from any other previous recording and to leave some gap between two consecutive recordings, silence may be recorded before and after the actual sound. The durations of the initial and final silence have to be specified by the user, for which the display is

ENTER INITIAL & FINAL DURATIONS IN SECS

G GRAPH on printer.

Make sure that printer is on.

The immediate question following the command acceptance is

WHICH PARAMETER?

for the parameter specified in response, its graph against time will be plotted on the printer.

A wrong parameter name will produce the error message

INVALID PARAMETER

and the program will wait for a new parameter name.

V Graphic display on VDU.

As in G, first the parameter name has to be entered when asked for. Next, the computer will ask

WHICH SEGMENT?

As in case of D, due to the screen size limitation only 35 frames can be plotted at a time. Thus the whole range of frames is divided in segments of 35 frames each. A segment number should be smaller than [INT (TOTAL No. of frames/35) + 1].

A segment display will remain on screen as long as the user does not press any key; as soon as any key is pressed, the screen will clear and the program will again ask for segment no. When the user does not want to display any other segment, a 0 should be entered for segment no. to eliminate the V command.

E EXIT from the program.
This will halt the program.

3. Some Guidelines for Data Manipulation

i) The number of points should be greater than one and less than or equal to 256.

ii) The value of FI (frame interval) should be less than 40.

iii) Each parameter value should lie in its range. A table of parameter limit values is given below.

iv) The value of PD (the point duration) should be a multiple of FI (the frame interval).

v) For a point lasting one frame only, PD can have value 0 or be equal to FI.

vi) For a parameter which has to remain steady during the duration of a point and whose value at the next point is different, a minus sign should be assigned to its value.

vii) As - 0 can't be distinguished from 0, a parameter which has to remain steady at value 0 should be assigned the value -1.

viii) If the last point has a duration value greater than FI all non-zero parameter values should be assigned a minus sign and 0 need not be replaced by -1. If PD is 0 or FI, the minus sign is optional.

ix) The value of PD should never have a minus sign.

4. Precauations and Suggestions

1) Always take out the program tape after loading the program into memory.

2) When reading data from tape, give the correct filename.

3) If the data is being saved on tape for later use, keep a record of the tape number and filename. If any datafiles are not required any more, delete their names from tape directory so that they can be used.

4) Do not end a command by pressing RETURN. On the other hand always terminate data values by RETURN.

5) If data for a large number of points has to be entered it may be entered in sections. The command O followed by NO makes this possible to begin with. Enter data for a limited number, save it and append new points.

6) It is better for large sections of data to be saved before proceeding further, as one may accidentally destroy data.

7) If one wants to use same tape for saving more than one datafile it is possible only if the next data is entered and saved in the same operation. When the first datafile has been saved, do not rewind the tape and do not remove it.

8) Whenever the command G has been executed and the printer has to be used for any other purpose, switch the printer off and then on again.

5. Error Messages

An error message may be generated by the program or by the system. If it is generated by the program, the execution will not be terminated, whereas if the system has generated it, the program execution may or may not terminate.

Program messages:

1) INVALID COMMAND. The message is displayed when the user tries to enter a command which does not belong to the program commands described in the previous section. It may be displayed during execution of command P, which has to be terminated by X, otherwise even the valid commands will be treated as invalid.

2) ILLEGAL SOURCE. This message can be generated if input data device entered is other than K or T. The program will ask for data source again.

3) NO. OF PTS OUT OF RANGE, DATA INVALID. If a datafile has not been created by the program but has been created independently and the number of points does not lie in the

range $2 \leq N \leq 256$, the above message is displayed, the data is rejected and the program waits for a new command.

4) INVALID No. If the number entered for NUMBER OF POINTS IS < 2 or > 256, the number is rejected with the above message and a new number is asked for.

5) POINT OUT OF RANGE. This error message can be enforced when one tries to modify a point number which does not exist in memory.

6) INVALID PARAMETER. If a parameter name other than one of the ten program parameters AV, . . . , PD is entered, the above message is displayed. It can also be generated by the commands V or G, (which are not applicable to PD) if PD is typed in.

7) SEGMENT DOES NOT EXIST. The message may be generated by either D or V command if a segment not present in memory is called for. If the message is displayed while trying to see the first data segment, it implies that if there was any data, it has been destroyed somehow.

8) TOTAL NO. EXCEEDS 256, ENTER AGAIN. Under the command O, when the data points are being appended, if the sum of the number of points in memory and the number to be added exceeds 256, the above message is generated. A new smaller value should be entered.

9) INVALID COMMAND, FRAMES NOT CREATED. The commands, P, R, G and V can not be executed until the intermediate frame values have been calculated by command C. If any of these commands is entered before C has been executed, the above message will be displayed and the program waits for a new command. C should be executed first, and then the

required command entered when '*' is displayed.

10) MODIFY PD POINT NUMBER. The message can be generated while either entering the data or modifying PD value for any point and the value of PD entered is not a multiple of FI. The next line will display the two nearest values which are multiples of FI, as follows.

SHOULD BE EITHER 'L1' OR 'L2'

where L1 and L2 are values less than and greater than the actual value of PD. The user has to decide the new value and enter it.

The above message can also be generated while shifting PD values for all points using the command O or reading a datafile created independently. In such a case, the message will be repeated for all points having an invalid PD value and the user must not be down the point numbers to modify each illegal value.

11) MODIFY PD. This message is generated by command C if the illegal values of PD for the points displayed in message 10 (generated by command O or tape input) have not been modified. User must detect and modify error values.

Error messages generated by system

1) RE DO FROM START. The message is generated if the type of data entered does not match with the type required. e.g. if numeric data is entered for a parameter name or vice versa. The program execution does not terminate and entering the correct value will continue the program.

2) ILLEGAL QUANTITY ERROR IN 1340. The message can be generated while executing command C, if either a parameter value is out of its range or a last point parameter value does not have a minus sign which it should have.

The program execution will be terminated and to enter the program again, type in

```
GOTO 180
```

Then check and modify the error values.

For any other message, consult the

PET USER MANUAL

TABLE OF PARAMETERS' LIMIT VALUES

Parameter	Low Value	High Value
AV	0	255
FO	74	464
F1	175	1452
F2	527	4356
F3	1704	5508
AH	0	255
AF	0	255
FF	1705	14160
AN	0	255

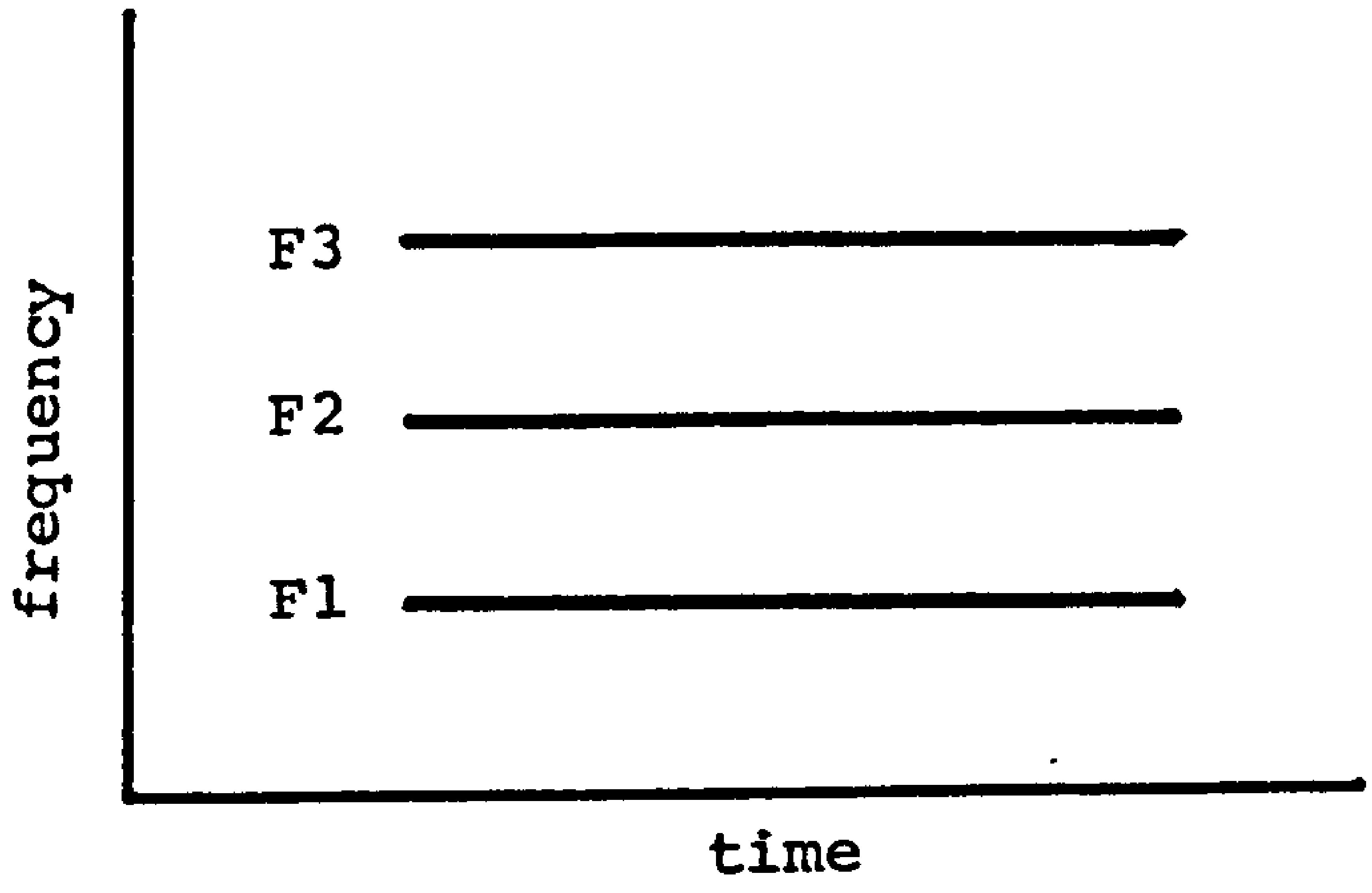


Fig. (1)

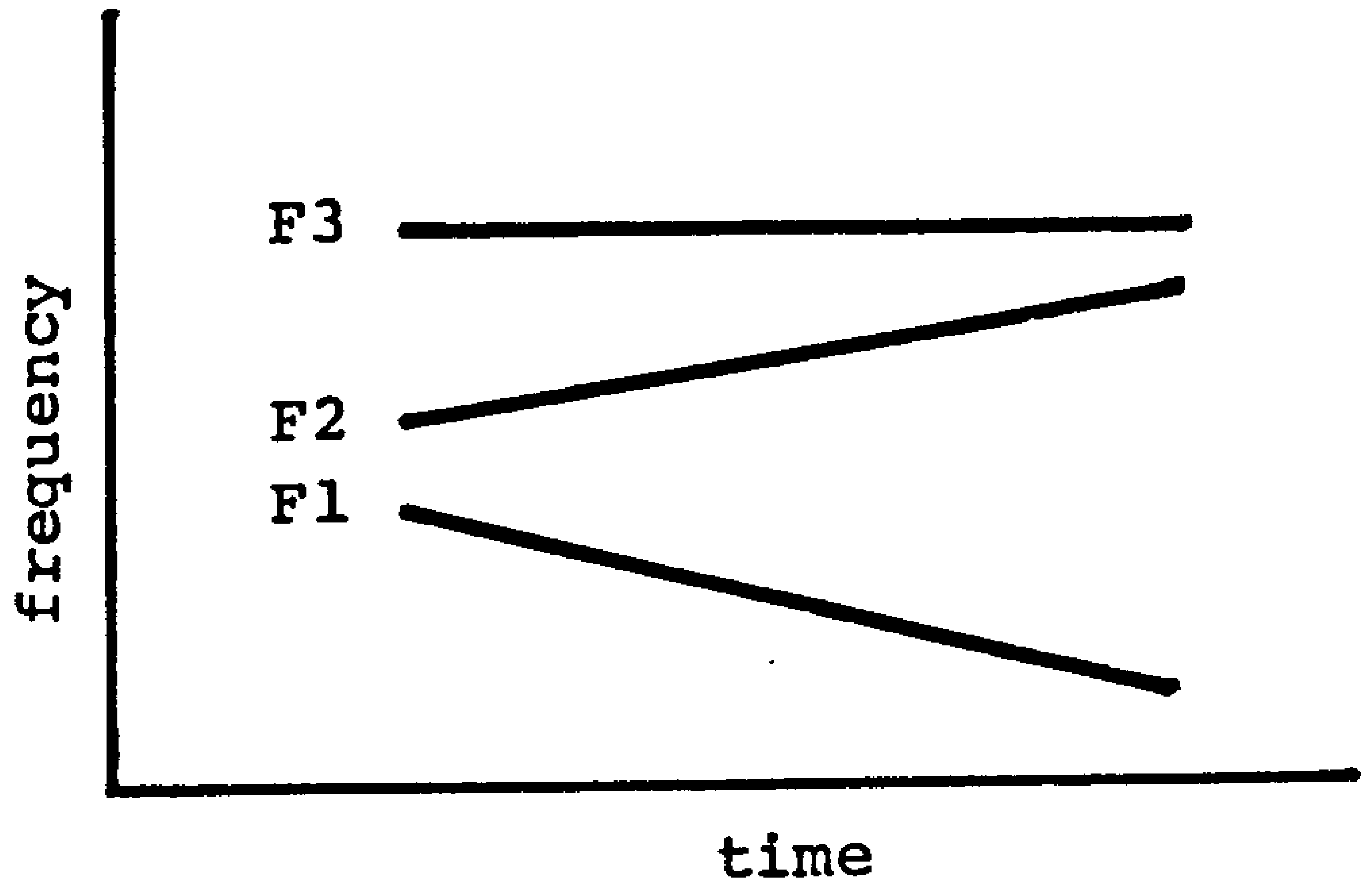


Fig. (2)

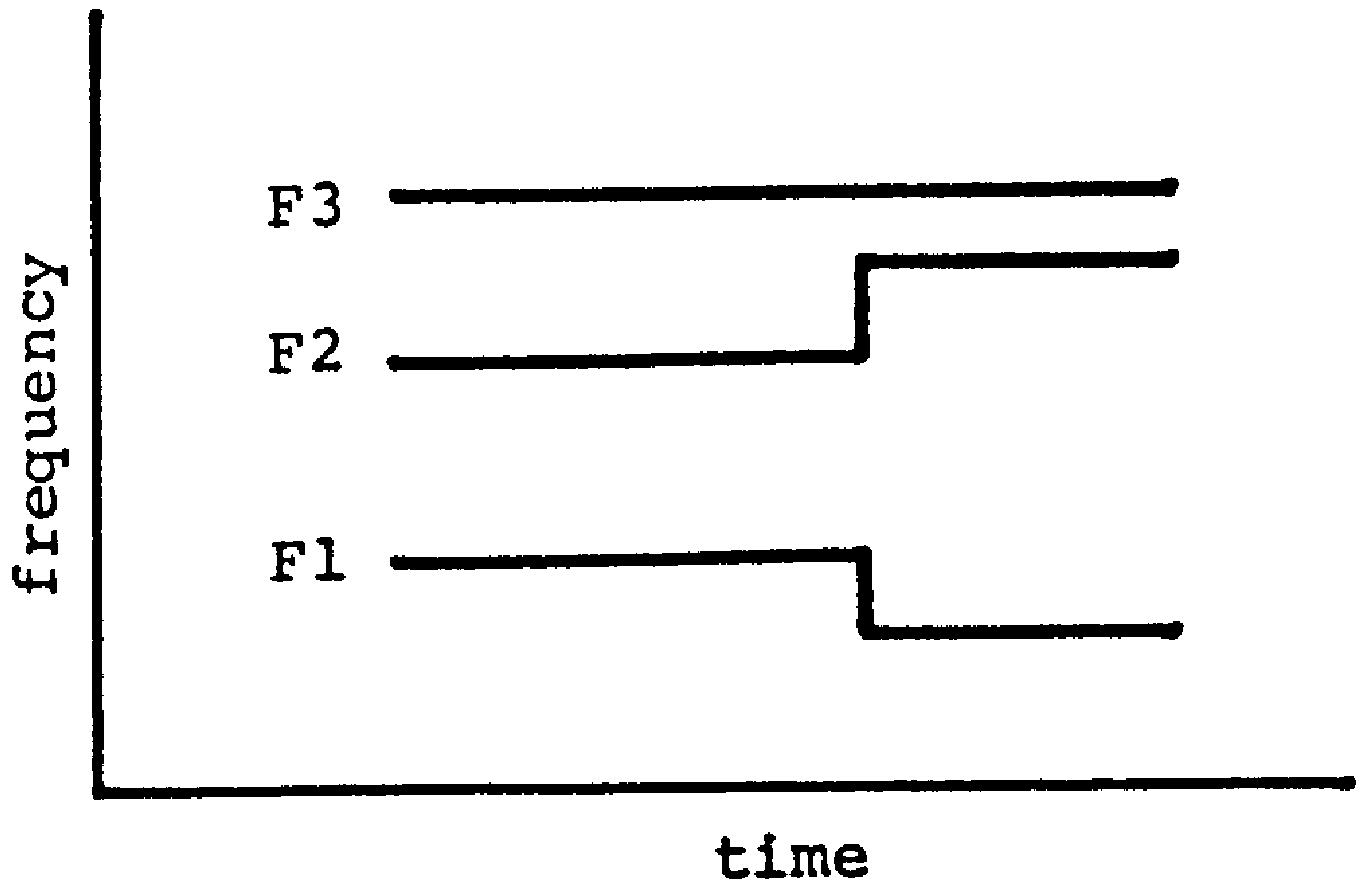


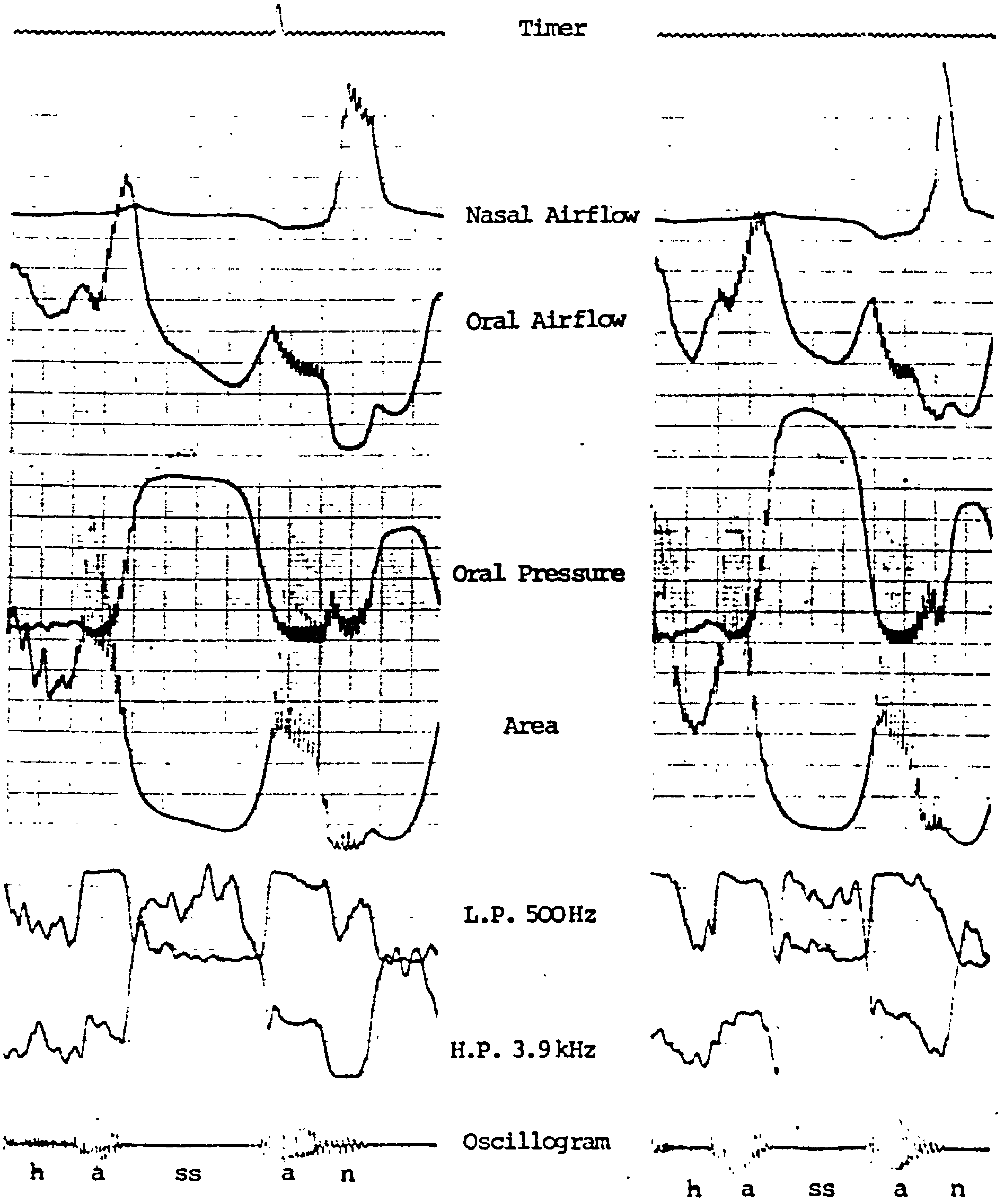
Fig. (3)

Appendix (2)

Photograph showing the instrumentation employed in the synthetic speech experiments.

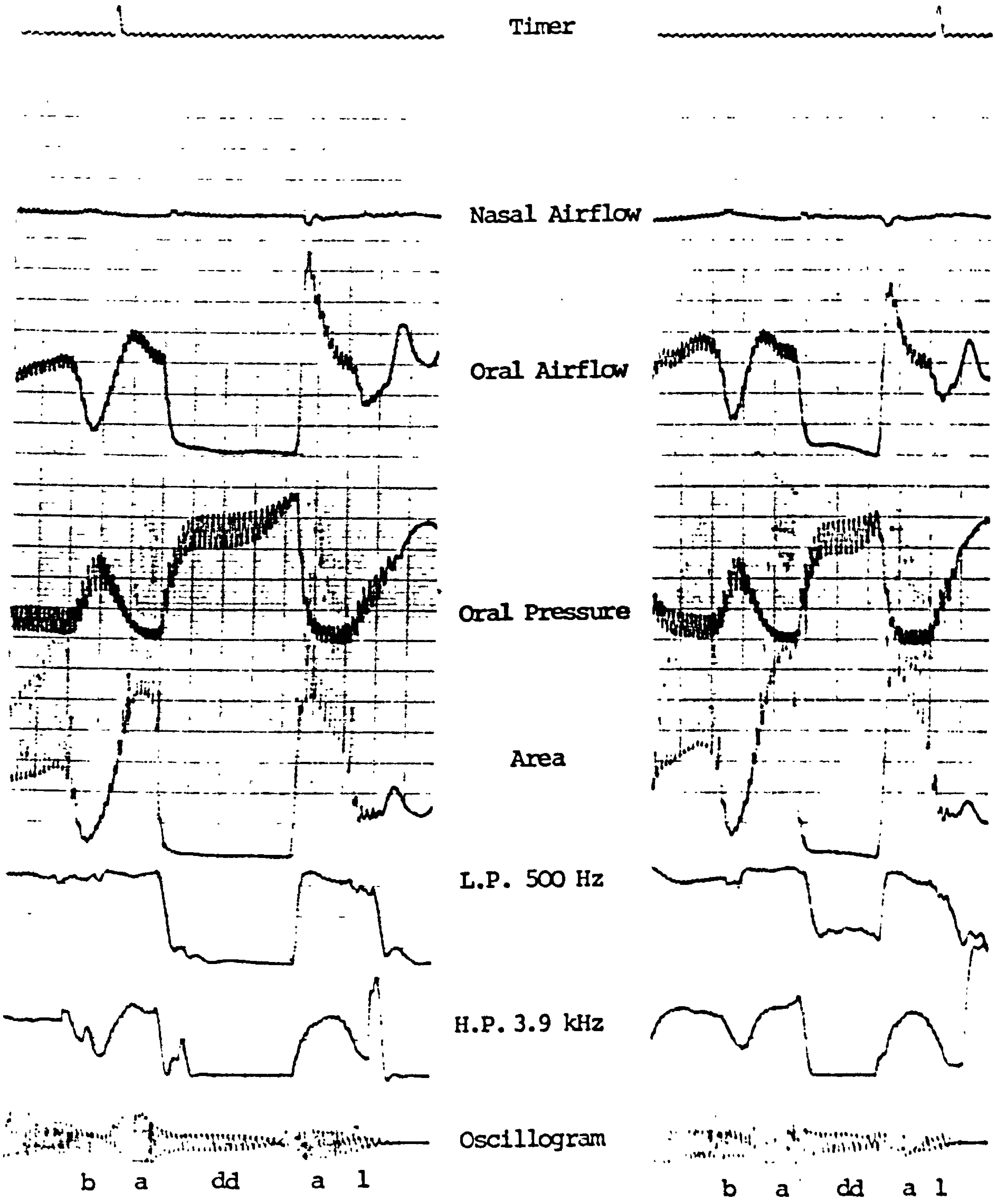

FERROGRAPH
PET

Appendix (3)



Two mingograms of the word /ħassan/ in the carrier sentence /'kitbi 'ħassan 'sit mar'raat/ as spoken by subject GG. Left : produced at slow speaking rate; right : produced at fast speaking rate.

Appendix (4)



Two mingograms of the word /baddal/ in the carrier sentence /'kitbi 'baddal 'sit mar'raat/ as spoken by subject AM. Left : produced at slow speaking rate; right : produced at fast speaking rate.

Appendix (5)

The SPSS file used for ANOVA in the main experiment on the articulation and aerodynamics of speech production.

```
PAGESIZE          NOEJECT
RUN NAME          1983 EXPERIMENT
VARIABLE LIST     SPKR,RATE,POSN,TYPE,SINGEM,T1,T4,T2,AC,PO,UO,ACDURC,
                  ACDURV
INPUT MEDIUM      DISK
N OF CASES        192
INPUT FORMAT      FREEFIELD
COMPUTE           ACVCRATIO=ACDURC/ACDURV
COMPUTE           ARVCRATIO=T2/T1
VAR LABELS        SPKR,SPEAKER/
                  RATE,SPEAKING RATE/
                  POSN,POSITION/
                  TYPE,TYPE OF CONSONANT/
                  SINGEM,SINGLE VS GEMINATE/
                  T1,ARTICULATORY TIMING OF PRECEDING VOWEL/
                  T4,ARTICULATORY TIMING BETWEEN AC MIN OF H AND AC MIN
                  OF S AND SS/
                  T2,ARTICULATORY DURATION OF CONONANT/
                  AC,MINIMUM CROSS SECTIONAL AREA OF SUPRAGLOTTAL
                  CONSTRICTION/
                  PO,ORAL PRESSURE/
                  UO,VOLUME AIRFLOW RATE/
                  ACDURC,ACOUSTIC DURATION OF CONSONANT/
                  ACDURV,ACOUSTIC TIMING OF PRECEDING VOWEL/
                  ACVCRATIO,ACOUSTIC DURATION OF CONSONANT DIVIDED BY
                  ACOUSTIC TIMING OF PRECEDING VOWEL/
                  ARVCRATTIO,ARTICULATORY DURATION OF CONSONANT DIVIDED
                  BY ARTICULATORY TIMING OF PRECEDING VOWEL
VALUE LABELS      SPKR (1) GG (2) AM/
                  RATE (1) SLOW (2) FAST/
                  POSN (1) INITIAL (2) MEDIAL/
                  TYPE (1) FRICATIVE (2) PLOSIVE/
                  SINGEM (1) SINGLE (2) GEMINATE
ANOVA             T2 BY SPKR(1,2) RATE(1,2) POSN(1,2) TYPE(1,2) SINGEM(1,2)
OPTIONS           1,4
```

Appendix (6)

The SPSS file used for the t-tests in the main experiment on the articulation and aerodynamics of speech production.

```
PAGESIZE         NOEJECT
RUN NAME         1983 EXPERIMENT
VARIABLE LIST    SPKR,RATE,POSN,TYPE,SINGEM,T1,T4,T2,AC,PO,UO,
                 ACDURC,ACDURV
INPUT MEDIUM     DISK
N OF CASES       192
INPUT FORMAT     FREEFIELD
COMPUTE          ACVCRATIO=ACDURC/ACDURV
COMPUTE          ARVCRATIO=T2/T1
VAR LABELS       SPKR,SPEAKER/
                 RATE,SPEAKING RATE/
                 POSN,POSITION/
                 TYPE,TYPE OF CONSONANT/
                 SINGEM,SINGLE VS GEMINATE/
                 T1,ARTICULATORY TIMING OF PRECEDING VOWEL/
                 T4,ARTICULATORY TIMING BETWEEN AC MIN OF H AND AC
                 MIN OF S AND SS/
                 T2,ARTICULATORY DURATION OF CONSONANT/
                 AC,MINIMUM CROSS SECTIONAL AREA OF SUPRAGLOTTAL
                 CONSTRICTION/
                 PO,ORAL PRESSURE/
                 UO,VOLUME AIRFLOW RATE/
                 ACDURC,ACOUSTIC DURATION OF CONSONANT/
                 ACDURV,ACOUSTIC TIMING OF PRECEDING VOWEL/
                 ACVCRATIO,ACOUSTIC DURATION OF CONSONANT DIVIDED
                 BY ACOUSTIC TIMING OF PRECEDING VOWEL/
                 ARVCRATIO,ARTICULATORY DURATION OF CONSONANT DIVIDED
                 BY ARTICULATORY TIMING OF PRECEDING VOWEL
VALUE LABELS     SPKR (1) GG (2) AM/
                 RATE (1) SLOW (2) FAST/
                 POSN (1) INITIAL (2) MEDIAL/
                 TYPE (1) FRICATIVE (2) PLOSIVE/
                 SINGEM (1) SINGLE (2) GEMINATE
SELECT IF        (SPKR EQ 1 AND POSN EQ 2 AND TYPE EQ 1 AND SINGEM EQ 2)
T-TEST           GROUPS=RATE/VARIABLES=T2
OPTIONS          1,4
```

Appendix (7)

Table (1A)

Arvcratio values[1] obtained from dividing the articulatory duration of the consonant (T_2) by the articulatory timing of the preceding vowel (T_1). See tables 6.39, 6.41, 6.43 and 6.45.

T_2	T_1	Arvcratio
54	93.0	0.58
54	85.3	0.63
47	93.0	0.51
47	93.0	0.51
54	93.0	0.58
54	93.0	0.58
47	93.0	0.51
54	77.5	0.70
47	77.5	0.61
54	93.0	0.58
47	93.0	0.51
47	85.3	0.55
171	124.0	1.38
155	100.8	1.54
140	131.8	1.06
124	131.8	0.94
155	116.3	1.33
132	124.0	1.06
116	93.0	1.25
109	85.3	1.28
101	77.5	1.30
109	93.0	1.17
109	85.3	1.25
101	100.8	1.00

---

[1] The values are rounded up to the second decimal place.

T_2	T_1	Arvcratio
39	85.3	0.46
39	85.3	0.46
47	77.5	0.61
39	77.5	0.50
39	85.3	0.46
39	85.3	0.46
39	69.8	0.56
31	85.3	0.36
31	69.8	0.44
31	69.8	0.44
31	69.8	0.44
39	77.5	0.50
155	85.3	1.82
163	85.3	1.91
163	85.3	1.91
147	77.5	1.90
155	77.5	2.00
140	77.5	1.81
85	77.5	1.10
109	77.5	1.41
93	69.8	1.33
109	77.5	1.41
101	77.5	1.30
109	77.5	1.41
47	85.3	0.55
54	108.5	0.50
70	62.0	1.13
62	77.5	0.80
62	77.5	0.80
62	93.0	0.67
70	62.0	1.13
62	93.0	0.67
47	62.0	0.76
39	77.5	0.50

T_2	T_1	Arvcratio
54	93.0	0.58
54	73.6	0.73
140	77.5	1.81
124	85.3	1.45
194	93.0	2.10
140	89.1	1.57
163	93.0	1.75
147	77.5	1.90
109	85.3	1.28
132	69.8	1.89
101	77.5	1.30
116	85.3	1.36
116	85.3	1.36
124	93.0	1.33
54	89.1	0.61
47	77.5	0.61
54	69.8	0.77
62	85.3	0.73
54	77.5	0.70
54	85.3	0.63
47	77.5	0.61
47	77.5	0.61
39	77.5	0.50
39	69.8	0.56
39	81.4	0.48
39	77.5	0.50
171	100.8	1.70
140	93.0	1.51
147	85.3	1.72
140	89.1	1.57
155	93.0	1.67
140	77.5	1.81
124	89.1	1.39

T_2	T_1	Arvcratio
124	77.5	1.60
124	85.3	1.45
124	85.3	1.45
124	77.5	1.60
132	77.5	1.70

Table (2A)

Acvcratio values[1] obtained from dividing the acoustic duration of the consonant (t_2) by the acoustic timing of the preceding vowel (t_1). See tables 6.39, 6.41, 6.43 and 6.45.

t_2	t_1	Acvcratio
85.3	69.8	1.22
92.0	69.8	1.33
77.5	77.5	1.00
77.5	69.8	1.11
85.3	62.0	1.38
77.4	62.0	1.25
62.0	69.8	0.89
77.5	62.0	1.25
69.8	69.8	1.00
77.5	62.0	1.25
77.5	69.8	1.11
69.8	62.0	1.13
193.8	85.3	2.27
162.8	85.3	1.91
178.3	69.8	2.55
170.5	69.8	2.44
178.3	69.8	2.55
170.5	62.0	2.75
131.8	69.8	1.89
131.8	62.0	2.13
116.3	62.0	1.88
131.8	77.5	1.70
124.0	69.8	1.78
116.3	69.8	1.67

---

[1] The values are rounded up to the second decimal place.

t_2	t_1	Acvcratio
31.0	85.3	0.36
31.0	93.0	0.33
38.8	85.3	0.45
38.8	85.3	0.45
31.0	85.3	0.36
31.0	85.3	0.36
31.0	77.5	0.40
31.0	77.5	0.40
31.0	62.0	0.50
31.0	69.8	0.44
31.0	69.8	0.44
38.8	69.8	0.56
139.5	93.0	1.50
147.3	100.8	1.46
147.3	85.3	1.73
139.5	85.3	1.64
139.5	85.3	1.64
131.8	77.5	1.70
77.5	85.3	0.91
100.8	69.8	1.44
93.0	69.8	1.33
93.0	85.3	1.09
100.8	77.5	1.30
93.0	77.5	1.20
62.0	69.8	0.89
77.5	100.8	0.77
108.5	54.3	2.00
96.9	62.0	1.56
77.5	77.5	1.00
85.3	100.8	0.85
69.8	65.9	1.06
65.9	89.1	0.74
73.6	62.0	1.19

t_2	t_1	Acvcratio
62.0	69.8	0.89
93.0	77.5	1.20
69.8	62.0	1.13
155.5	77.5	2.01
139.5	81.4	1.71
193.8	73.6	2.63
170.5	77.5	2.20
189.9	77.5	2.45
166.7	69.8	2.39
124.0	73.6	1.68
155.0	69.8	2.22
139.5	62.0	2.25
147.3	69.8	2.11
143.4	69.8	2.05
139.5	73.6	1.90
50.4	96.9	0.52
46.5	85.3	0.55
50.4	77.5	0.65
54.3	93.0	0.68
46.5	77.5	0.60
46.5	100.8	0.46
42.6	77.5	0.55
46.5	93.0	0.50
42.6	77.5	0.55
31.0	93.0	0.33
38.8	85.3	0.45
34.9	77.5	0.45
170.5	100.8	1.69
135.7	93.0	1.46
147.3	85.3	1.73
135.7	96.9	1.40
155.0	93.0	1.67

t_2	t_1	Acvcratio
139.5	96.9	1.44
131.8	93.0	1.42
131.8	93.0	1.42
131.8	85.3	1.55
124.0	93.0	1.33
127.9	77.5	1.65
131.8	77.5	1.70

BIBLIOGRAPHY

BIBLIOGRAPHY

Abercrombie, D., (1957). 'Direct Palatography, 'Zeitschrift Für Phonetik, vol. 10, pp.21-25, reprinted in Abercrombie, D., (1965). Studies in Phonetics and Linguistics. London: Oxford University Press, pp.125-130.

Abercrombie, D., (1965). Studies in Phonetics and Linguistics. London: Oxford University Press; especially pp.120-124.

Abercrombie, D., (1967). Elements of General Phonetics. Edinburgh: Edinburgh University Press.

Ainsworth, W.A., (1971). 'Perception of Synthesized Isolated Vowels and h-d Words as a Function of Fundamental Frequency, 'J. Acoust. Soc. Am., vol.49, pp.1323-1324.

Ainsworth, W.A. and Miller, J.B., (1972). 'Identification of Synthetic Isolated Vowels and Vowels in h-d Context,' Acoustica, vol. 27, pp.278-282.

Al-Ani, S.H., (1967). 'An Acoustical and Physiological Investigation of the Arabic /ξ/, 'in Readings in Arabic Linguistics, edited by Al-Ani, S.H. (1978); Indiana University Linguistics Club, Bloomington, Indiana, pp.89-101.

Al-Ani, S.H., (1970). Arabic Phonology: An Acoustical and Phonological Investigation. The Hague : Mouton.

Al-Ani, S.H., (1978). Readings in Arabic Linguistics (ed.). Indiana University Linguistics Club, Bloomington, Indiana.

Al-Ani, S.H. and May, D.R., (1973). 'The Phonological Structure of the Syllable in Arabic,' in Readings in Arabic Linguistics, edited by Al-Ani, S.H. (1978); Indiana University Linguistics Club, Bloomington, Indiana, pp.113-125.

Al-Anṭāki, M., (n.d.) ʔal wadʒiiz fi fiqh il luɣa. "A Summary of Philology." Damascus.

Allan, G.D., (1970). 'Structure of Timing in Speech Production,' J.Acoust.Soc.Am., vol.47, p.58 (Abstract).

Al-Tajir, M.A., (1982). Language and Linguistic Origins in Bahrain. Library of Arabic Linguistics, Monograph No. 5. London: Kegan Paul International.

Anīs, I., (1971). *?al ?aṣwaat il luɣawijja*. "The Linguistic Sounds". Cairo: Anglo-Egyptian Bookshop.

Anthony, J.K., (1954). 'A New Method of Investigating The Tongue Positions of Consonants,' *Science Technologists' Bulletin*, Oct.-Nov., pp.2-5.

Aoki, P.K., (1981). *Gemination in Japanese*. Unpublished Ph.D. Dissertation, University of Washington.

Arkebauer, H.J.,(1964). *A Study of Intraoral Air Pressure Associated With Production of Selected Consonants*. Unpublished Ph.D. Dissertation, State University of Iowa.

Arkebauer, H.J., Hixon, T.J. and Hardy, J.C., (1967). 'Peak Intraoral Air Pressures During Speech,' *J.Speech Hear.Res.*, vol.10, pp.196-208.

Ashby, M.G., (1979). Review of A. Malécot's 'Contribution à l' étude de la force d'articulation en français (The Hague: Mouton), *JIPA*, vol.9, No.2, pp.78-80.

Balasubramanian, T., (1982). 'Intervocalic Double Nasal and Lateral Consonant Articulations in Tamil,' *J.Phon.*, vol.10, pp.99-104.

Barkai, M., (1974). 'On Duration and Spirantization in Biblical Hebrew,' *Linguistic Inquiry*, vol.5, pp.456-459.

Bastian, J., Eimas, P.D. and Liberman, A.M., (1961), 'Identification and Discrimination of a Phonemic Contrast Induced by Silent Interval,' *J.Acoust.Soc.Am.*, vol.55, p.456 (Abstract).

Beeston, A.F.L., (1970). *The Arabic Language Today*. London: Hutchinson University Library.

Belasco, S., (1953). 'The Influence of Force of Articulation of Consonants on Vowel Duration,' *J. Acoust. Soc. Am.*, vol. 25, pp.1015-1016.

Benhallam, A., (1979). 'Geminates: A New Approach,' *Proceedings of the Mid-American Linguistics Conference*, edited by Cooley, R., Barnes, M. and Dunn, J., Morman: The University of Oklahoma, pp.212-221.

Benhallam, A., (1980). *Syllable Structures and Rule Types in Arabic*. Unpublished Ph.D. Dissertation, University of Florida.

Bernthal, J.E. and Beukelman, D.R., (1978). 'Intraoral Air Pressure During the Production of /p/ by Children, Youths, and Adults,' *J.Speech Hear.Res.*, vol.21, pp.361-371.

Birkeland, H., (1954). Stress Patterns in Arabic. Oslo: Jacob Dybwad.

Bishr, K.M., (1970). ξilm il luγa lξaam. "General Linguistics." Cairo: daar il maξaarif bimaṣr.

Black, J.W., (1950). 'The Pressure Component in the Production of Consonants,' J.Speech Hear.Dis., vol.15, pp.207-210.

Blanc, H.,(1952). Studies in North Palestinian Arabic. Jerusalem: Central Press.

Blanc, H., (1964). Communal Dialects in Baghdad. Cambridge: Harvard University Press.

Blanc, H., (1967). 'The Sonorous vs. Muffled Distinction in Old Arab Phonology,' in Readings in Arabic Linguistics, edited by Al-Ani, S.H. (1978); Indiana University Linguistics Club, Bloomington, Indiana, pp.127-146.

Bloch, B., (1950). 'Studies in Colloquial Japanese IV: Phonemics,' Language, vol.26, pp.86-125, reprinted in Reading in Linguistics I, edited by Martin Joos (1966); Chicago: The University of Chicago Press, 4th ed., pp.329-348.

Blumstein, S.E. and Stevens, K.N., (1979). 'Acoustic Invariance in Speech Production: Evidence from Measurements of the Spectral Characteristics of Stop Consonants,' J.Acoust.Soc.Am., vol.66, pp.1001-1017.

Blumstein, S.E. and Stevens, K.N., (1980). 'Perceptual Invariance and Onset Spectra for Stop Consonants in Different Vowel Environments,' J.Acoust.Soc.Am., vol.67, pp.648-662.

Blumstein, S.E. and Stevens, K.N., (1981). 'Phonetic Features and Acoustic Invariance in Speech,' Cognition, vol.10, pp.25-32.

Bonnot, J-F.P.,(1979). 'Étude Experimentale De Certains Aspects De La Gemination Et De L'Emphase En Arabe,' Proceedings of the Ninth International Congress of Phonetic Sciences, vol.1, Copenhagen, p.180.

Borden, G.J. and Harris, K.S., (1980). Speech Science Primer: Physiology, Acoustics, and Perception of Speech. Baltimore: The Williams and Wilkins Company.

Brown, W.S., (1969). An Investigation of Intraoral Pressures During Production of Selected Consonants. Unpublished Ph.D. Dissertation, State University of New York at Buffalo.

Brown, W.S. and McGlone, R.E., (1969). 'Relation of Intra-oral Air Pressure to Oral Cavity Size,' Folia Phoniatrica, vol.21, pp.321-331.

Brown, W.S. and Shearer, W.M., (1970). 'Constancy of Intra-oral Air Pressure Related to Integrated Pressure Time Measurements,' Folia Phoniatrica, vol. 22, pp.49-57.

Brown, W.S., McGlone, R.E. and Proffit, M.R., (1973). 'Relationship of Lingual and Intraoral Air Pressures During Syllable Production, J.Speech Hear.Res., vol.16, pp.141-151.

Burns, E.M. and Ward, W.D., (1974). 'Categorical Perception of Musical Intervals,' J.Acoust.Soc.Am., vol.55, p.456 (Abstract).

Butcher, A. and Weiher, E., (1976). 'An Electropalatographic Investigation of Coarticulation in VCV Sequences,' J.Phon., vol.4, pp.59-74.

Cantineau, J., (1960a). Études De Linguistique Arabe. Paris: Librairie C. Klincksieck.

Cantineau, J., (1960b). Cours De Phonétique Arabe. Paris: Librairie C. Klincksieck.

Carney, A.E. and Widin, G.P., (1976). 'Acoustic Discrimination Within Phonetic Categories,' J.Acoust.Soc.Am., vol.59, S24 (Abstract).

Carney, A.E., Widin, G.P. and Viemeister, N.F., (1977). 'Non-Categorical Perception of Stop Consonants Differing in VOT,' J.Acoust.Soc.Am., vol.62, pp.961-970.

Carnochan, J., (1962). 'Gemination in Hausa, 'Studies in Linguistic Analysis (Special Volume of the Philological Society), Oxford: Basil Blackwell, pp.149-181.

Catford, J.C., (1977). Fundamental Problems in Phonetics. Edinburgh: Edinburgh University Press.

Centmayer, K., (1975). 'Interrelations of Vowel Perception and Linguistic Context,' in Auditory Analysis and Perception of Speech, edited by Fant, G. and Tatham, M.A.A.; London: Academic Press, pp.143-152.

Chen, M., (1970). 'Vowel Length Variation as a Function of the Voicing of the Consonant Environment,' Phonetica, vol.22, pp.129-159.

Chomsky, N. and Halle, M., (1968). The Sound Pattern of English. New York: Harper and Row.

Cole, O.F., (1872). 'On the Production of Articulate Sound (Speech),' *Transactions of the Odontological Society*, vol.4, 2nd series, pp.110-123.

Cole, R.A. and Scott, B., (1974). 'Towards a Theory of Speech Perception,' *Psychological Review*, vol.81, pp.348-374.

Cooper, W.E., (1976). 'Syntactic Control of Timing in Speech Production,' *J.Phon.*, vol.4, pp.151-171.

Cooper, F.S., Delattre, P.C., Liberman, A.M., Borst, J.M. and Gerstman, L.J., (1952). 'Some Experiments on the Perception of Synthetic Speech Sounds,' *J.Acoust.Soc.Am.*, vol.24, pp.597-606.

Cowan, W., (1958). *An Introduction to Modern Literary Arabic*. London: Cambridge University Press.

Cowan, W., (1966). 'Review of Communal Dialects in Baghdad,' by Blanc, H. (1964), *Language*, vol.42, pp.694-699.

Cowell, M.W., (1964). *A Reference Grammar of Syrian Arabic*. Washington D.C. : Georgetown University Press.

Creelman, C.D., (1962). 'Human Discrimination of Auditory Duration,' *J.Acoust.Soc.Am.*, vol.34, pp.582-593.

Crystal, D., (1969). *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press.

Crystal, D., (1980). *A First Dictionary of Linguistics and Phonetics*. London: André Deutsch Limited.

Cutting, J.E. and Rosner, B.S., (1974). 'Categories and Boundaries in Speech and Music,' *Perception and Psychophysics*, vol.16, pp.564-570.

Delattre, P., (1962). 'Some Factors on Vowel Duration and their Cross-Linguistic Validity,' *J.Acoust.Soc.Am.*, vol.34, pp.1141-1143.

Delattre, P., (1971). 'Consonant Gemination in Four Languages: An Acoustic, Perceptual and Radiographic Study,' *IRAL*, vol.9, part 1, pp.31-52; part 2, pp.97-113.

Delattre, P., Liberman, A.M. and Cooper, F.S., (1951). 'vwajɛl sɛ̃tetik a dø fɔrmɑ̃t e vwajɛl kardinal,' *Le Maître Phonétique*, No.96, pp.30-36.

Delattre, P.C., Liberman, A.M. and Cooper, F.S., (1955). 'Acoustic Loci and Transitional Cues for Consonants,' *J.Acoust.Soc.Am.*, vol.27, pp.769-773.

Denes, P., (1955). 'Effect of Duration on the Perception of Voicing,' *J.Acoust.Soc.Am.*, vol.27, pp.761-764.

Denes, P.B. and Pinson, E.N., (1963). *The Speech Chain: The Physics and Biology of Spoken Language*. New York: Anchor Books.

Eimas, P.D., (1975). 'Auditory and Phonetic Coding of the Cues for Speech: Discrimination of the (r-l) Distinction by Young Infants,' *Perception and Psychophysics*, vol.18, pp.341-347.

Ellis, S.M. and Weismer, G., (1978). 'More on Temporal Compensation,' *J.Acoust.Soc.Am.*, vol. 63, pp.S54.

El-Saaran, M.H.A., (1951). *A Critical Study of the Phonetic Observations of the Arab Grammarians*. Unpublished Ph.D. Thesis, University of London.

El-Saaran, M.H.A., (1962). *ʕilm il luɣa : muqaddima lil qaari? il ʕarabi*. "Linguistics: An Introduction to the Arabic Reader." Cairo: daar il maʕaarif bimaṣr.

Erwin, W.M., (1963). *A Short Reference Grammar of Iraqi Arabic*. Washington D.C. :Georgetown University Press.

Erwin, W.M., (1969). *A Basic Course in Iraqi Arabic*. Washington D.C. : Georgetown University Press.

Fant, G., (1960). *Acoustic Theory of Speech Production*. The Hague: Mouton.

Fant, G., (1968). 'Analysis and Synthesis of Speech Processes,' in *Manual of Phonetics*, edited by Malmberg, B.; Amsterdam: North-Holland Publishing Company, pp.173-277.

Fant, G., (1973). *Speech Sounds and Features*. The MIT Press.

Fant, G. and Tatham, M.A.A., (1975). *Auditory Analysis and Perception of Speech*. (eds.) London: Academic Press.

Ferguson, C.A., (1956). 'Review of Stress Patterns in Arabic,' by Birkeland, H., (1954), *Language*, vol.32, pp.384-387.

Ferguson, C.A., (1975). 'Two Problems in Arabic Phonology,' *Word*, vol.13, pp.460-478.

Ferguson, G.A., (1959). *Statistical Analysis in Psychology and Education*. New York : McGraw-Hill Book Company, Inc.

Ferrero, F.E., Pelamatti, G.M. and Vagges, K., (1982). 'Continuous and Categorical Perception of a Fricative-Affricate Continuum,' *J.Phon.*, vol.10, pp.231-244.

Fintoft, K., (1961). 'The Duration of Some Norwegian Speech Sounds,' Phonetica, vol.7, pp.19-39.

Firth, J.R., (1948). 'Word-Palatograms and Articulation,' Bulletin of the School of Oriental and African Studies, vol.12, pp.857-864; reprinted in Firth, J.R. (1957). Papers in Linguistics 1934-1951. London: Oxford University Press, pp.146-155.

Fischer-Jørgensen, E., (1964). 'Sound Duration and Place of Articulation,' Zeitschrift für Phonetik, vol.17, pp.175-207.

Fischer-Jørgensen, E., (1972). 'Perceptual Studies of Danish Stop Consonants,' ARIPUC, vol.6, pp.75-176.

Fitch, H.L., (1981). 'Distinguishing Temporal Information for Speaking Rate from Temporal Information for Intervocalic Stop Consonant Voicing,' Haskins Laboratories, Status Report on Speech Research, SR-65, pp.1-32.

Fledge, J.E. and Port, R., (1981). 'Cross-Language Phonetic Interference: Arabic to English,' Language and Speech, vol.24, pp.125-146.

Forbes, D., (1863). A Grammar of the Arabic Language. London: W.H. Allen and Company Limited.

Fowler, C.A., (1977). Timing Control in Speech Production. Indiana University Linguistics Club, Bloomington, Indiana.

Fowler, C.A., (1980). 'Coarticulation and Theories of Extrinsic Timing,' J.Phon., vol.8, pp.113-133.

Fry, D.B., (1955). 'Duration and Intensity as Physical Correlates of Linguistic Stress,' J.Acoust.Soc.Am., vol.27, pp.765-768.

Fry, D.B., (1964). 'Experimental Evidence for the Phoneme,' in In Honour of Daniel Jones, edited by Abercrombie, D., Fry, D.B., MacCarthy, P.A.D., Scott, N.C. and Trim, J.L.M.; London: Longman, Green and Company Limited, pp.59-72.

Fry, D.B., (1973). 'Linguistic Theory and Experimental Research, 'in Phonetics in Linguistics : A Book of Readings, edited by Jones, W.E. and Laver, J.; London: Longman Group Limited, pp.66-87.

Fry, D.B., (1979). The Physics of Speech. Cambridge: Cambridge University Press.

Fry, D.B., Abramson, A.S., Eimas, P.D. and Liberman, A.M., (1962). 'The Identification and Discrimination of Synthetic Vowels,' *Language and Speech*, vol.5, pp.171-189.

Fujimura, O., (1983). 'Remarks on Speech Synthesis,' in *Abstracts of the Tenth International Congress of Phonetic Sciences*, edited by Cohen, A. and van den Broecke, M.P.R.; Foris Publications, Dordrecht-Holland, pp.113-119.

Fujisaki, H. and Sudo, H., (1971); 'Synthesis by Rule of Prosodic Features of Connected Japanese,' *Proceedings of the Seventh International Congress of Acoustics*, vol.3, pp.133-136.

Fujisaki, H., Nakamura, K. and Imoto, T., (1975). 'Auditory Perception of Duration of Speech and Non-Speech Stimuli,' in *Auditory Analysis and Perception of Speech*, edited by Fant, G. and Tatham, M.A.A.; London: Academic Press, pp.197-219.

Gairdner, W.H.T., (1925). *The Phonetics of Arabic*. London: Oxford University Press.

Gairdner, W.H.T., (1935). 'The Arab Phoneticians on the Consonants and Vowels,' in *Readings in Arabic Linguistics*, edited by Al-Ani, S.H. (1978); Indiana University Linguistics Club, Bloomington, Indiana, pp.187-202.

Gaitenby, J., (1965). 'The Elastic Word,' *Haskins Laboratories, Status Report on Speech Research*, SR-2, pp.3.1-3.12.

Gandhi, J. and Roach, P., (1980). *Instruction Manual - SPS256: A Program for Speech Synthesis with a PET Computer and a Computalker Speech Synthesiser*. Department of Linguistics and Phonetics, University of Leeds.

Gay, T., (1968). 'Effect of Speaking Rate on Diphthong Formant Movements,' *J.Acoust.Soc.Am.*, vol.44, pp.1570-1573.

Gay, T., (1978). 'Effect of Speaking Rate on Vowel Formant Movements,' *J.Acoust.Soc.Am.*, vol.63, pp.223-230.

Gay, T., (1981a). 'Temporal and Spatial Properties of Articulatory Movements: Evidence for Minimum Spreading Across and Maximum Effects Within Syllable Boundaries,' in *The Cognitive Representation of Speech*, edited by Myers, T., Laver, J. and Anderson, J.; North-Holland Publishing Company, pp.133-138.

Gay, T., (1981b). 'Mechanisms in the Control of Speech Rate,' *Phonetica*, vol.38, pp.148-158.

Gay, T. and Hirose, H., (1973). 'Effect of Speaking Rate on Labial Consonant Production,' Phonetica, vol.27, pp.44-56.

Gay, T., Ushijima, T., Hirose, H. and Cooper, F.S., (1974). 'Effect of Speaking Rate on Labial Consonant-Vowel Articulation,' J.Phon., vol.2, pp.47-63.

Gay, T. and Ushijima, T., (1975). 'Effect of Speaking Rate on Stop Consonant-Vowel Articulation,' in Speech Communication, Proceedings of Speech Communication Seminar, Stockholm, edited by Fant, G.; Uppsala: Almqvist and Wiksell, pp.205-208.

Gerstman, L.J., Liberman, A.M., Delattre, P.C. and Cooper, F.S., (1954). 'Rate and Duration of Changes in Formant Frequency as Cues for the Identification of Speech Sounds,' J.Acoust.Soc.Am., vol.26, p.952 (Abstract).

Ghalib, G.B.M., (1977). The Intonation of Colloquial Iraqi Spoken Arabic. Unpublished M.Phil. Thesis, Department of Phonetics, University of Leeds.

Gimson, A.C., (1970). An Introduction to the Pronunciation of English. (2nd ed.). London: Edward Arnold.

Gleason, H.A., (1955). An Introduction to Descriptive Linguistics. New York: Henry Hold and Company.

Goldman-Eisler, F., (1968). Psycholinguistics: Experiments in Spontaneous Speech. New York: Academic Press.

Guilford, J.P. and Fruchter, B., (1978). Fundamental Statistics in Psychology and Education. (6th ed.). Tokyo: McGraw-Hill Kogakusha, Limited.

Haag, W.K., (1977). 'Oral Air Pressure and Airflow in German Stop Consonants,' Phonetica, vol.34, pp.31-39.

Haggard, M., (1973). 'Abbreviation of Consonants in English Pre- and Post-Vocalic Clusters,' J.Phon., vol.1, pp.9-24.

Hall, R.A., (1964). Introductory Linguistics. New York: Chilton Books.

Han, M.S., (1965). 'The Feature of Duration in Japanese,' J.Phon.Soc.Jap., (Study and Sound), Vol.10, pp.65-80.

Hanna, S.A. and Greis, N., (1972). Beginning Arabic. Leiden : E.J. Hill.

Hardcastle, W., (1968). 'Dynamic Palatography,' Work in Progress, Department of Phonetics and Linguistics, University of Edinburgh, vol.2, pp.53-57.

Hardcastle, W., (1969). 'A System of Dynamic Palatography,' *Work in Progress*, Department of Phonetics and Linguistics, University of Edinburgh, vol.3, pp.47-52.

Hardcastle, W., (1970). 'Electropalatography in Speech Research,' *Occasional Papers*, Language Centre, University of Essex, vol.9, pp.54-64.

Hardcastle, W., (1971). *Electropalatography in the Investigation of Some Physiological Aspects of Speech Production*. Unpublished Ph.D. Thesis, Department of Phonetics and Linguistics, University of Edinburgh.

Hardcastle, W., (1972). 'The Use of Electropalatography in Phonetic Research,' *Phonetica*, vol.25, pp.197-215.

Hardcastle, W., (1978). 'Palatography,' *Manuscript*, pp.1-6.

Hardcastle, W. and Roach, P.J., (1976). 'A Computer System for the Processing of Electropalatography and Other Articulatory Data,' *Proceedings of Phonetics Symposium, Occasional Papers*, No.17, Language Centre, University of Essex.

Hardy, J.C., (1965). 'Air Flow and Air Pressure Studies,' *Proceedings Conference: Communicative Problems in Cleft Palate, ASHA Reports*, vol.1, part 1, pp.141-152.

Harrell, R.S., (1957). *The Phonology of Colloquial Egyptian Arabic*. New York : American Council of Learned Societies.

Harrell, R.S., (1962). 'Consonant, Vowel and Syllable in Moroccan Arabic, '*Proceedings of the Fourth International Congress of Phonetic Sciences*. The Hague: Mouton.

Harris, K.S., (1954). 'Cues for the Identification of the Fricatives of American English,' *J.Acoust. Soc.Am.*, vol. 26, p.952 (Abstract).

Harris, K.S., Hoffman, H.S., Liberman, A.M., Delattre, P.C. and Cooper, F.S.,(1958). 'Effect of Third-Formant Transition on the Perception of the Voiced Stop Consonants,' *J.Acoust.Soc.Am.*, vol.30, pp.122-126.

Ḥassaan, T., (1955). *manaahidʒ il baħθ fi l luɣa*. "Research Methods in Language." Cairo: Anglo-Egyptian Bookshop.

Hassan, Z.M., (1981). *An Experimental Study of Vowel Duration in Iraqi Spoken Arabic*. Unpublished Ph.D. Thesis, Department of Linguistics and Phonetics, University of Leeds.

Hegedüs, L., (1959). 'Beitrag zur Frage der Geminaten,' Zeitschrift für Phonetik, vol.12, pp.68-106.

Heinz, J.M. and Stevens, K.N., (1961). 'On the Properties of Voiceless Fricatives,' J.Acoust. Soc.Am., vol.33, pp.589-596.

Henry, F.M., (1948). 'Discrimination of the Duration of a Sound,' J.Exper.Psych., vol.38, pp.734-743.

Hixon, T.J., (1966). 'Turbulent Noise Sources of Speech,' Folia Phoniatrica, vol.18, pp.168-182.

Hockett, C.F., (1955). A Manual of Phonology. Baltimore: Waverly Press, Inc.

Hopkin, G.B. and McEwen, J.D., (1955). 'Speech Defects and Malocclusion: A Palatographic Investigation,' The Dental Practitioner, vol.6, pp.123-131.

House, A.S., (1961). 'On Vowel Duration in English,' J.Acoust.Soc.Am., vol.33, pp.1174-1178.

House, A.S. and Fairbanks, G., (1953). 'The Influence of Consonant Environment Upon the Secondary Acoustical Characteristics of Vowels,' J.Acoust.Soc.Am., Vol.25, pp.105-113.

Howell, M.S., (1883). A Grammar of the Classical Arabic Language. Allahabad.

Huggins, A.W.F., (1972). 'Just Noticeable Differences for Segment Duration in Natural Speech,' J.Acoust.Soc.Am., vol.51, pp.1270-1290.

Huitson, A., (1966). The Analysis of Variance. Charles Griffin.

Hyman, L., (1975). Phonology: Theory and Analysis. New York: Holt, Rinehart and Winston.

Ingham, B., (1975). A Study of the Phonology and Morphology of the Verbal Piece of the Arabic of Khūzistan. Unpublished Ph.D. Thesis, University of London.

International Phonetic Association, The Principles of the (1949). London: University College London. Reprinted in 1970.

Isshiki, N., (1965). 'Vocal Intensity and Airflow Rate,' Folia Phoniatrica, vol.17, pp.92-104.

Isshiki, N. and Ringel, R., (1964). 'Airflow During the Production of Selected Consonants,' J.Speech Hear Res., vol.7, pp.233-244.

Jaeger, J.J., (1978). 'Speech Aerodynamics and Phonological Universals,' Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society, pp.311-329.

Johnstone, T.M., (1967). Eastern Arabian Dialect Studies. London: Oxford University Press.

Jones, D., (1956). The Pronunciation of English. (4th ed.) Cambridge: Cambridge University Press.

Jones, D., (1960). An Outline of English Phonetics. (9th ed.) Cambridge: Cambridge University Press.

Jones, D., (1967). The Phoneme : Its Nature and Use. (3rd ed.) Cambridge: Cambridge University Press.

Josselyn, F., (1901). 'Études Expérimentales de la Phonétique Italienne,' La Parole, vol.11, pp.227-252.

Karnell, M.P. and Willis, C.R., (1982a). 'The Effect of Vowel Context on Consonantal Intraoral Air Pressure,' Folia Phoniatrica, vol.34, pp.1-8.

Karnell, M.P. and Willis, C.R., (1982b). 'The Reliability of the Kay Agnellograph Pressure Transducer in the Study of Consonantal Intraoral Air Pressure,' Folia Phoniatrica, vol.34, pp.53-56.

Kelly, J., Anthony, J.K. and Uldall, E., (1962). 'Tempo and Transitions,' Proceedings of Speech Communication Seminar, R.I.T., Stockholm, paper B9.

Kenstowicz, M.J. and Kisseberth, C.W., (1973). Issues in Phonological Theory. (eds.) The Hague: Mouton.

Kenstowicz, M. and Pyle, C., (1973). 'On the Phonological Integrity of Geminate Clusters,' in Issues in Phonological Theory : Proceedings of the Urbana Conference on Phonology, edited by Kenstowicz, M. and Kisseberth, C.W.; The Hague : Mouton, pp.27-43.

Kent, R.D. and Moll, K.L., (1969). 'Vocal-tract Characteristics of the Stop Cognates,' J.Acoust.Soc.Am., vol.46, pp.1549-1555.

Kent, R.D. and Moll, K.L., (1972a). 'Tongue Body Articulation During Vowel and Diphthong Gestures,' Folia Phoniatrica, vol.24, pp.286-300.

Kent, R.D. and Moll, K.L., (1972b). 'Cinefluorographic Analysis of Selected Lingual Consonants,' J.Speech Hear.Res., vol.15, pp.453-473.

Kent, R., Carney, P. and Severeid, L., (1974). 'Velar Movement and Timing: Evaluation of a Model for Binary Control,' J.Speech Hear.Res., vol.17, pp.470-488.

Kent, R.D. and Moll, K.L., (1975). 'Articulatory Timing in Selected Consonant Sequences,' Brain and Language, vol.2, pp.304-323.

Kent, R.D. and Minifie, F.D., (1977). 'Coarticulation in Recent Speech Production Models,' J.Phon., vol.5, pp.115-133.

Kingsley, N.W., (1880). Oral Deformities. London: H.K. Lewis.

Klatt, D.H., (1967). 'Articulatory Activities and Air Flow During the Production of Fricative Consonants,' Quarterly Progress Report, MIT, vol.15, pp.257-260.

Klatt, D.H., (1973). 'Interaction Between Two Factors that Influence Vowel Duration,' J.Acoust.Soc.Am., vol.54, pp.1102-1104.

Klatt, D.H., (1974). 'The Duration of [s] in English Words,' J.Speech Hear.Res., vol.17, pp.51-63.

Klatt, D., (1975). 'Vowel Lengthening is Syntactically Determined in a Connected Discourse,' J.Phon., vol.3, pp.129-140.

Klatt, D., (1976). 'The Linguistic Uses of Segment Duration in English: Acoustic and Perceptual Evidence,' J.Acoust.Soc.Am., vol.59, pp.1208-1221.

Klatt, D.H., Stevens, K.N. and Mead, J., (1968). 'Studies of Articulatory Activity and Airflow During Speech,' in Sound Production in Man, Annals of the New York Academy of Sciences, edited by Bouhuys, A., vol.155, art.1, pp.42-55.

Kozhevnikov, V.A. and Chistovich, L.A., (1965). Speech Articulation and Perception. Moscow-Leningrad: Nauka. Translated into English by U.S. Department of Commerce, Joint Publication Research Service, Washington D.C.

Kuhl, P.K. and Miller, J.D., (1975). 'Speech Perception by the Chinchilla: Voiced-Voiceless Distinction in Alveolar Plosive Consonants,' Science, vol. 190, pp.69-72.

Kunze, L.H., (1964). 'Evaluation of Methods of Estimating Subglottal Air Pressure,' J.Speech Hear.Res., vol.7, pp.151-164.

Kurylowicz, J., (1972). Studies in Semitic Grammar and Metrics. Wroclaw : Zaklad Narodowy Imienia Ossolinskich Wyawrictwo, Polskiej Akademii Nauk.

Kydd, W.L. and Belt, D.A., (1964). 'Continuous Palatography,' J.Speech Hear.Dis., vol.29, pp.489-492.

Ladefoged, P., (1957). 'Use of Palatography,' J.Speech Hear. Dis., vol.22, pp.764-774.

Ladefoged, P., (1963). 'Some Physiological Parameters in Speech,' Language and Speech, vol. 6, pp.109-119.

Ladefoged, P., (1971). Preliminaries to Linguistics Phonetics. Chicago: The University of Chicago Press.

Ladefoged, P., (1975). A Course in phonetics. New York : Harcourt Brace Jovenovich, Inc.

Ladefoged, P. and McKinney, N.P., (1963). 'Loudness, Sound Pressure, and Subglottal Pressure in Speech,' J.Acoustic.Soc.Am., vol.35, pp.454-460.

Lane, H., (1965). 'The Motor Theory of Speech Perception: A Critical Review,' Psychological Review, vol.72, pp.275-309.

Laver, J., (1968). 'Phonetics and the Brain,' Work in Progress, Department of Linguistics and Phonetics, University of Edinburgh, vol.2, pp.63-75.

Laver, J., (1970). 'The Production of Speech,' in New Horizons in Linguistics, edited by Lyons, J.; Penguin Books, pp.53-75.

Leben, W.R., (1980). 'A Metrical Analysis of Length,' Linguistic Inquiry, vol.11, pp.497-509.

Lehiste, I., (1960). 'Segmental and Syllabic Quantity in Estonian,' American Studies in Uralic Linguistics. Bloomington, Indiana, Indiana University Press, pp.21-82.

Lehiste, I., (1966). Consonant Quantity and Phonological Units in Estonian. Indiana University Publications, Uralic and Altaic Series, vol.65, Bloomington, Indiana, Indiana University.

Lehiste, I., (1970). Suprasegmentals. The MIT Press.

Lehiste, I., (1971). 'Temporal Organization of Spoken Language,' in Form and Substance, Phonetic and Linguistic Papers Presented to Eli Fischer-Jørgensen, edited by Hammerich, L.L., Jakobson, R. and Zwirner, E.; Copenhagen : Akademisk, pp.159-169.

Lehiste, I., (1972a). 'Temporal Compensation in a Quantity Language,' Proceedings of the Seventh International Congress of Phonetic Sciences. The Hague : Mouton, pp.929-937.

Lehiste, I., (1972b). 'The Timing of Utterances and Linguistic Boundaries,' J. Acoust.Soc. Am., vol.51, pp.2018-2024.

Lehiste, I., (1973). 'Rhythmic Units and Syntactic Units in Production and Perception,' J.Acoust.Soc.Am., vol.54, pp.1228-1234.

Lehiste, I. and Peterson, G.E., (1959). 'Vowel Amplitude and Phonemic Stress in American English,' J.Acoust. Soc.Am., vol.31, pp.428-435.

Lehiste, I. and Peterson, G.E., (1961). 'Transitions, Glides and Diphthongs,' J.Acoust.Soc.Am., vol.33, pp.268-277.

Lehiste, I., Morton, K. and Tatham, M.A.A., (1973). 'An Instrumental Study of Consonant Gemination,' J.Phon., vol.1, pp.131-148.

Liberman, A.M., (1957). 'Some Results of Research on Speech Perception,' J.Acoust.Soc.Am., vol.29, pp.117-123.

Liberman, A.M., (1970). 'Some Characteristics of Perception in the Speech Mode,' in Perception and Its Disorders, edited by Hamburg, D.A.; New York: Williams and Wilkins.

Liberman, A.M., Delattre, P.C. and Cooper, F.S., (1952). 'The Role of Selected Stimulus-Variables in the Perception of the Voiced Stop Consonants,' Am.J.Psych., vol.65, pp.497-516.

Liberman, A.M., Delattre, P.C., Cooper, F.S. and Gerstman, L.J., (1954). 'The Role of Consonant - Vowel Transitions in the Perception of the Stop and Nasal Consonants,' Psychological Monographs, vol.68 (8, Whole No.379).

Liberman, A.M., Harris, K.S., Hoffman, H.S. and Griffith, B.C., (1957). 'The Discrimination of Speech Sounds Within and Across Phoneme Boundaries,' J.Exper.Psych., vol.54, pp.358-368.

Liberman, A.M., Delattre, P.C. and Cooper, F.S., (1958). 'Some Cues for the Distinction Between Voiced and Voiceless Stops in Initial Position,' Language and Speech, vol.1, pp.153-167.

Liberman, A.M., Ingemann, F., Lisker, L., Delattre, P.C., Cooper, F.S., (1959). 'Minimal Rules for Synthesizing Speech,' J.Acoust.Soc.Am., vol.31, pp.1490-1499.

Liberman, A.M., Harris, K.S., Eimas, P.D., Lisker, L. and Bastian, J, (1961a). 'An Effect of Learning on Speech Perception: The Discrimination of Durations of Silence with and without Phonemic Significance,' Language and Speech, vol.4, pp.175-195.

Liberman, A.M., Harris, K.S., Kinney, J.A. and Lane, H., (1961b). 'The Discrimination of Relative Onset Time of the Components of Certain Speech and Non-speech Patterns,' J.Exper.Psych., vol.61, pp.379-388.

Liberman, A.M., Stevens, K.N. and Öhman, S., (1963). 'Identification and Discrimination of Synthetic Speech Sounds,' Report No. STL-QPSR-2, Royal Institute of Technology, Stockholm.

Liberman, A.M., Cooper, F.S., Harris, K.S. and MacNeilage, P.F., (1963). 'A Motor Theory of Speech Perception,' Proceedings of Speech Communication Seminar, Royal Institute of Technology, Stockholm .

Liberman, A.M., Cooper, F.S., Shankweiler, D.P. and Studdert-Kennedy, M., (1967). 'Perception of the Speech Code,' Psychological Review, vol.74, pp.431-461.

Lindblom, B., (1963). 'Spectrographic Study of Vowel Reduction,' J.Acoust.Soc.Am., vol.35, pp.1773-1781.

Lindblom, B., (1968). 'Temporal Organization of Syllable Production,' Speech Transmission Laboratory Quarterly Progress Status Report, Publication No.2, Royal Institute of Technology, Stockholm.

Lindblom, B., (1970). 'On the Temporal Organization of Syllabic Processes,' J.Acoust.Soc.Am., vol.48, p.111.

Lindblom, B. and Rapp, K., (1973). 'Some Temporal Regularities of Spoken Swedish,' Institute of Linguistics, University of Stockholm, Publication No. 21.

Lindquist, E.F., (1940). Statistical Analysis in Educational Research. Boston : Houghton Mifflin Company.

Lisker, L., (1957). 'Closure Duration and the Voiced-Voiceless Distinction in English,' Language, vol.33, pp.257-267.

Lisker, L., (1970). 'Supraglottal Air Pressure in the Production of English Stops,' Language and Speech, vol.13, pp.215-230.

Lisker, L. and Abramson, A.S., (1964). 'A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements,' Word, vol.20, pp.384-422.

Lubker, J.F., (1969). 'Velopharyngeal Orifice Area: A Replication of Analog Experimentation,' J.Speech Hear. Res., vol.12, pp.218-222.

Lubker, J.F. and Moll, K.L., (1965). 'Simultaneous Oral-Nasal Air Flow Measurements and Cinefluorographic Observations During Speech Production,' Cleft Palate Journal, vol.2, pp.257-272.

MacCarthy, P.A.D., (1950). English Pronunciation. (4th ed.). Cambridge : Heffer.

Mack, M. and Blumstein, S.E., (1983). 'Further Evidence of Acoustic Invariance in Speech Production: The Stop-Glide Contrast,' J.Acoust.Soc.Am., vol.73, pp.1739-1750.

MacNeilage, P.F. and DeClerk, J.T., (1969). 'Cinefluorographic Study of Speaking Rate,' J.Acoust.Soc.Am., vol.45, p.308.

Malécot, A., (1955). 'An Experimental Study of Force of Articulation,' Studia Linguistica, vol.9, pp.35-44.

Malécot, A., (1966). 'The Effectiveness of Intraoral Air-Pressure-Pulse Parameters in Distinguishing Between Stop Cognates,' Phonetica, vol.2, pp.65-81.

Malécot, A., (1970). 'The Lenis-Fortis Opposition: Its Physiological Parameters,' J.Acoust.Soc.Am., vol.47, pp.1588-1592.

Martinet, A., (1964). Elements of General Linguistics. London: Faber and Faber. Translated into English by E. Palmer.

McGlone, R.E. and Proffit, W.R., (1972). 'Dynamic Palatography of Selected Syllables,' Proceedings of the Seventh International Congress of Phonetic Sciences, part 1, pp.375-378.

McKay, G.R., (1980). 'Medial Stop Gemination in Rembarrnga: A Spectrographic Study,' J.Phon., vol.8, pp.343-352.

Meddis, R., (1973). Elementary Analysis of Variance for the Behavioural Sciences. London: McGraw-Hill Book Company (UK) Limited.

Miller, G.A. and Nicely, P.E., (1955). 'An Analysis of Perceptual Confusions Among Some English Consonants,' J.Acoust.Soc.Am., vol.27, pp.338-352.

Miller, J.D., Wier, C.C., Pastore, R.E., Kelly,W.J. and Dooling, R.J., (1976). 'Discrimination and Labelling of Noise-Buzz Sequences with Varying Noise-Lead Times: An Example of Categorical Perception,' J.Acoust.Soc.Am., vol.60, pp.410-417.

Miller, J.L. and Eimas, P.D., (1977). 'Studies on the Perception of Place and Manner of Articulation: A Comparison of the Labial-Alveolar and Nasal-Stop Distinctions,' J.Acoust.Soc.Am., vol.61, pp.835-845.

Miller, J.L. and Liberman, A.M., (1979). 'Some Effects of Later-Occurring Information on the Perception of Stop Consonant and Semivowel,' Perception and Psychophysics, vol.25, pp.457-465.

Miller, J. and Baer, T., (1983). 'Some Effects of Speaking Rate on the Production of /b/ and /w/,' J.Acoust. Soc.Am., vol.73, pp.1751-1755.

Miller, S., (1975). Experimental Design and Statistics. London: Methuen and Company Limited.

Mitchell, T.F., (1956). An Introduction to Egyptian Colloquial Arabic. London : Oxford University Press.

Mitchell, T.F., (1962). 'Long Consonants in Phonology and Phonetics,' Studies in Linguistic Analysis (Special Volume of the Philological Society). Oxford: Basil Blackwell, pp.182-203.

Mitchell, T.F., (1975). Principles of Firthian Linguistics. London : Longman Group Limited.

Moore, B.C.J., (1977). Introduction to the Psychology of Hearing. The Macmillan Press Limited.

Morse, P.A. and Snowdon, C.T., (1975). 'An Investigation of Categorical Speech Discrimination by Rhesus Monkeys,' Perception and Psychophysics, vol.17, pp.9-16.

Mortimer, C., (1977). A Phonetic Study of Certain Grammatical Distinctions in English : Acoustic and Perceptual Aspects. Unpublished M.Phil. Thesis, Department of Phonetics, University of Leeds.

Moses, E.R., (1939). 'Palatography and Speech Improvement,' J.Speech Hear. Dis., vol.4, pp.103-114.

Muller, E.M. and Brown, W.S., (1980). 'Variations in the Supraglottal Air Pressure Waveform and their Articulatory Interpretation,' in Speech and Language: Advances in Basic Research and Practice, edited by Lass, N.J., vol.4, pp.317-389; London: Academic Press.

Murry, T. and Brown, W.S., (1979). 'Aerodynamic Interactions Associated with Voiced-Voiceless Stop Consonants,' Folia Phoniatrica, vol.31, pp.82-88.

Myers, J.L., (1966). Fundamentals of Experimental Design. Allyn and Bacon.

Nasr, R.T., (1959). 'The Predictability of Stress in Lebanese Arabic,' Phonetica, vol.4, pp.89-94.

Nasr, R.T., (1960). 'Phonemic Length in Lebanese Arabic,' Phonetica, vol.5, pp.209-211.

Netsell, R., (1969). "Subglottal and Intraoral Air Pressures During the Intervocalic Contrast of /t/ and /d/,' Phonetica, vol.20, pp.68-73.

Nooteboom, S.G., (1981). 'Speech Rate and Segmental Perception or the Role of Words in Phoneme Identification,' in The Cognitive Representation of Speech, edited by Myers, T., Laver, J. and Anderson, J.; Amsterdam : North-Holland Publishing Company, pp.143-150.

Nooteboom, S.G. and Slis, I.H., (1972). 'The Phonetic Feature of Vowel Length in Dutch,' Language and Speech, vol.15, pp.301-316.

Obrecht, D.H., (1965). 'Three Experiments in the Perception of Geminate Consonants in Arabic,' Language and Speech, vol.8, pp.31-41.

O'Connor, J.D., (1973). Phonetics. London: Penguin Books Limited.

O'Connor, J.D., Gerstman, L.J., Liberman, A.M., Delattre, P.C. and Cooper, F.S., (1957). 'Acoustic Cues for the Perception of Initial /w,j,r,l/ in English,' Word, vol.13, pp.25-43.

Odisho, E.Y., (1973). The Role of the Rear Section of the Vocal Tract in Baghdadi Arabic. Unpublished M.Phil. Thesis, Department of Phonetics, University of Leeds.

Odisho, E.Y., (1975). The Phonology and Phonetics of the Neo-Aramaic as Spoken by the Assyrian in Iraq. Unpublished Ph.D. Thesis, Department of Phonetics, University of Leeds.

Ohala, J.J., (1975). 'On the Temporal Regulation of Speech,' in Auditory Analysis and Perception of Speech, edited by Fant, G. and Tatham, M.A.A.; London : Academic Press, pp.431-453.

Oller, D.K., (1973). 'The Effect of Position in Utterance on Speech Segment Duration in English,' J.Acoust. Soc.Am., vol.54, pp.1235-1246.

Oussani, G., (1901). 'The Arabic Dialect of Baghdad,' Journal of the American Oriental Society, vol.22, pp.94-114.

Palmer, F.R., (1962). 'Gemination in Tigrinya,' Studies in Linguistic Analysis (Special Volume of the Philological Society), Oxford: Basil Blackwell, pp.139-149.

Pastore, R.E., (1976). 'Categorical Perception: A Critical Re-Evaluation,' in Hearing and Davis: Essays Honouring Hallowell Davis, edited by Hirsh, S.K., Eldredge, D.H., Hirsch, I.J. and Silverman, S.R.; Washington University Press, St. Louis, Missouri.

Pastore, R.E., Ahroon, W.A., Baffuto, K.J., Friedman, C., Puleo, J.S. and Fink, E.A., (1977). 'Common-Factor Model of Categorical Perception,' J. Exper. Psych. Human Perception and Performance, vol.3, pp.686-696.

Pei, M., (1966). Glossary of Linguistic Terminology. New York: Columbia University Press.

Pei, M. and Gaynor, F., (1954). Dictionary of Linguistics. London: Peter Owen Limited.

Perkell, J.S., (1965). 'Cineradiographic Studies of Speech: Implications of a Detailed Analysis of Certain Articulatory Movements,' Proceedings of the Fifth International Congress of Acoustics, vol. 1, paper A32, Université de Liège

Perkell, J.S., (1969). Physiology of Speech Prodction: Results and Implications of a Quantitative Cineradiographic Study. Research Monograph Series, No.53. Cambridge: The MIT Press.

Peterson, G.E., (1955). 'An Oral Communication Model,' Language, vol. 31, pp.414-427.

Peterson, G.E. and Lehiste, I., (1960). 'Duration of Syllable Nuclei in English,' J.Acoust.Soc.Am., vol.32, pp.693-703.

Pickett, J.M., (1980). The Sounds of Speech Communication. Baltimore: University Park.

Pickett, J.M. and Decker, L.R., (1960). 'Time Factors in Perception of a Double Consonant,' Language and Speech, vol.3, pp.11-17

Pike, K.L., (1947). Phonemics: A Technique for Reducing Language to Writing. Ann Arbor : University of Michigan Press.

Pisoni, D.B. and Lazarus, J.H., (1974). 'Categorical and Non-Categorical Modes of Speech Perception Along the Voicing Continuum,' J.Acoust.Soc.Am., vol.55, pp.328-333.

Port, R.F., Al-Ani, S. and Maeda, S., (1980). 'Temporal Compensation and Universal Phonetics,' Phonetica, vol.37, pp.235-252.

Prosek, R. and House, A.S., (1975). 'Intraoral Air Pressure as a Feedback Cue in Consonant Production,' J.Speech Hear.Res., vol.18, pp.133-147.

Rahim, A.J., (1980). The Phonology of Iraqi Spoken Arabic From a Functional Point of View . Unpublished Ph.D. Thesis, Department of Linguistics and Phonetics, University of Leeds.

Raphael, L.J., (1981). 'Durations and Contexts as Cues to Word-Final Cognate Opposition in English,' Phonetica, vol.38, pp.126-147.

Raphael, L.J., Dorman, M.F. and Liberman, A.M., (1980). 'On Defining the Vowel Duration that Cues Voicing in Final Position,' Language and Speech, vol. 23, pp.297-307.

Roach, H. and Roach, P., (1983). 'Automatic Segmentation and Labelling of Speech Sounds from Different Languages,' Working Papers in Linguistics and Phonetics, Leeds University, No. 1, pp.91-95.

Roach, P., (1980). 'Speech Synthesis - The PET That Talked Back,' Microcomputer Printout, vol.1, pp.9-11.

Roach, P., (1983). English Phonetics and Phonology : A Practical Course. Cambridge: Cambridge University Press.

Robins, R.H., (1980). General Linguistics : An Introductory Survey. (3rd ed.). London : Longman Group Limited.

Robson, C., (1973). Experiment, Design and Statistics in Psychology. Penguin Books.

Rodón, E., (1967). 'A Phonetic Realization of the Past Morpheme in English,' Proceedings of the Sixth International Congress of Phonetic Sciences, pp.773-774.

Roscoe, J.T., (1975). Fundamental Research Statistics for the Behavioural Sciences. (2nd ed.). New York : Holt, Rinehart and Winston, Inc.

Rothenberg, M., (1969). The Breath-Stream-Dynamics of Simple-Released-Plosive Production. Bibliotheca Phonetica, No. 6, Basel.

Rousselot, J., (1891). Les Modifications Phonétiques du Langage Étudiées dans le Patois d'une Famille de Cellefrouin. Paris.

Ruhm, H.B., Mencke, E.O., Milburn, B., Cooper, W.A. and Rose, D.E., (1966). 'Differential Sensitivity to Duration of Acoustic Signals,' J.Speech Hear.Res., vol.9, pp.371-384.

Sampson, G., (1973). 'Duration in Hebrew Consonants,' Linguistic Inquiry, vol.4, pp.101-104.

Scheffé, H.A., (1959). The Analysis of Variance. New York: Wiley.

Schwab, E., Sawusch, J. and Nusbaum, H.C., (1981). 'The Role of Second Formant Transitions in the Stop-Semivowel Distinction,' *Perception and Psychology*, vol.29, pp.121-128.

Schwartz, M.F., (1969). 'Influence of Vowel Environment Upon the Duration of /s/ and /ʃ/,' *J. Acoust.Soc. Am.*, vol.46, pp.480-481.

Schwartz, M.F., (1970). 'Duration of /s/ in /s/-Plosive Blends,' *J.Acoust. Soc.Am.*, vol.47, pp.1143-1144.

Scully, C., (1969). 'Problems in the Interpretation of Pressure and Air Flow Data in Speech,' *Phonetics Department Reports*, No.2, University of Leeds, pp.53-92.

Scully, C., (1971). 'A Comparison of /s/ and /z/ for an English Speaker,' *Language and Speech*, vol.14, pp.187-200.

Scully, C., (1975). 'A Synthesizer Study of Aerodynamic Factors in Speech Segment Durations,' in *Progress in Speech Communication. Proceedings of Speech Communication Seminar, Stockholm, 1974*, edited by Fant, G., vol.2, pp.227-234. Stockholm: Wiley.

Scully, C., (1979). 'Model Prediction and Real Speech: Fricative Dynamics,' in *Frontiers of Speech Communication Research*, edited by Lindblom, B. and Öhman, S.; London : Academic Press, pp.35-48.

Scully, C., (1982). *Departmental Handout*. Department of Linguistics and Phonetics, University of Leeds.

Scully, C., (1983). 'The Search for Cues (Acoustic and Articulatory Analysis),' *paper presented at 11^{e} ICA, Toulouse*, pp.1-15, (Manuscript).

Sharf, D.J., (1962). 'Duration of Post-Stress Intervocalic Stops and Preceding Vowels,' *Language and Speech*, vol.5, pp.26-30.

Scherer, R.C., Titze, I.R. and Curtis, T.S., (1983). 'Pressure-Flow Relationship in Two Models of the Larynx Having Rectangular Glottal Shapes,' *J.Acoust. Soc.Am.*, vol.73, pp.668-676.

Shockey, L., Gregorski, R. and Lehiste, I., (1971). 'Word-Unit Temporal Organization,' *J.Acoust. Soc. Am.*, vol.50, p.117.

Shohara, H.H. and Hanson, C., (1941). 'Palatography as an Aid to the Improvement of Articulatory Movements,' *J. Speech Hear. Dis.*, vol.6, pp.115-124.

Schouton, M.E.H., (1980). 'The Case Against a Speech Mode of Perception,' Acta Psychologia, vol.44, pp.71-98.

Siegel, S., (1956). Nonparametric Statistics for the Behavioural Sciences. Tokyo: McGraw-Hill Kogakusha Limited.

Sievers, E., (1893). Grundzüge der Phonetik. (4th ed.) Leipzig: Breitkopf and Hartel.

Slis, I.H., (1970). 'Articulatory Measurements on Voiced, Voiceless and Nasal Consonants,' Phonetica, vol.21, pp.193-210.

Slis, I.H., (1971). 'Articulatory Effect and Its Durational and Electromyographic Correlates,' Phonetica, vol.23, pp.171-188.

Slis, I.H. and Cohen, A., (1969). 'On the Complex Regulating the Voiced-Voiceless Distinction,' Language and Speech, vol.12, part 1, pp.80-102; part 2, pp.137-155.

Small, A.M. and Campbell, R.A., (1962). 'Temporal Differential Sensitivity for Auditory Stimuli,' Am.J.Psych. vol.75, pp.401-410.

Smith, B.E. and Weinberg, B., (1980). 'Prediction of Velopharyngeal Orifice Area: A Re-Examination of Model Experimentation,' Cleft Palate Journal, vol.17, pp.277-282.

Socin, A., (1895). Arabic Grammar. (2nd ed.). Translated from German by Kennedy, R.S.; London: Williams and Norgate.

Soli, S.D., (1982). 'Structure and Duration of Vowels together Specify Fricative Voicing,' J. Acoust. Soc.Am., vol.72, pp.366-378.

Stetson, R.H., (1951). Motor Phonetics. (2nd ed.). Amsterdam: North-Holland Publishing Company.

Stevens, K.N., (1975). 'The Potential Role of Property Detectors in the Perception of Consonants,' in Auditory Analysis and Perception of Speech, edited by Fant, G. and Tatham, M.A.A.; London: Academic Press, pp.303-330.

Stevens, K.N. and House, A.S., (1963). 'Perturbation of Vowel Articulations by Consonantal Context : An Acoustical Study,' J. Speech Hear. Res., vol.6, pp.111-128.

Stevens, K.N., Liberman, A.M., Studdert-Kennedy, M. and Öhman, S.E.G., (1969). 'Cross-Language Study of Vowel Perception,' Language and Speech, vol.12, pp.1-23.

Stevens, K.N. and House, A.S., (1972). 'Speech Perception,' in Foundations of Modern Auditory Theory, edited by Tobias, J.V.; vol.2, New York : Academic Press, pp.3-62.

Stevens, K.N. and Blumstein, S.E., (1978). 'Invariant Cues for Place of Articulation in Stop Consonants,' J.Acoust. Soc. Am., vol.64, pp.1358-1368.

Stevens, K.N. and Blumstein, S.E., (1981). 'The Search for Invariant Acoustic Correlates of Phonetic Features,' in Perspectives on the Study of Speech, edited by Eimas, P.D. and Miller, J.L.; Erlbaum, Hillsdale, pp.1-38.

Stevens, S.S. and Davis, H., (1938). Hearing : its Psychology and Physiology. New York : Wiley and Sons, Inc.

Stott, L.H., (1935). 'Time-Order Errors in the Discrimination of Short Tonal Durations,' J. Exper.Psych., vol.18, pp.741-766.

Strange, W., Jenkins, J.J. and Johnson, T.L., (1983). 'Dynamic Specification of Coarticulated Vowels,' J. Acoust. Soc. Am., vol.74, pp.695-705.

Strenger, F., (1968). 'Radiographic, Palatographic and Labiographic Methods in Phonetics,' in Manual of Phonetics, edited by Malmberg, B.; Amsterdam : North-Holland Publishing Company, pp.334-364.

Studdert-Kennedy, M., (1974). 'The Perception of Speech,' in Current Trends in Linguistics, edited by Sebeok, T.A., vol.12; The Hague : Mouton, pp.2349-2385.

Studdert-Kennedy, M., Liberman, A.M., Harris, K. and Cooper, F.S., (1970). 'The Motor Theory of Speech Perception : A Reply to Lane's Critical Review,' Psychological Review, vol.77, pp.234-249.

Subtelny, J.D., Worth, J.H. and Sakuda, M., (1966). 'Intraoral Pressure and Rate of Flow During Speech,' J. Speech Hear. Res., vol.9, pp.498-518.

Swadesh, M., (1937). 'The Phonemic Interpretation of Long Consonants,' Language, vol.13, pp.1-10.

Thananjayarajasingham, S., (1976a). 'Intervocalic Double Consonant Articulations in Ceylon Tamil : A Palatographic Study on Lengthening.' Language and Speech, vol.19, pp.75-79.

Thananjayarajasingham, S., (1976b). 'Duration of Intervocalic Stops in Ceylon Tamil,' Language and Speech, vol.19, pp.150-159.

Titze, I.R., (1980). 'Comments on the Myoelastic-Aerodynamic Theory of Phonation,' J. Speech Hear. Res., vol.23, pp.495-510.

Trubetzkoy, N.S., (1969). Principles of Phonology. Translated into English by Christiane A.M. Baltaxe. California: University of California Press Limited.

Tuller, B., Harris, K.S. and Kelso, J.A.S., (1981a). 'Articulatory Motor Events as a Function of Speaking Rate and Stress,' Haskins Laboratories, Status Report on Speech Research, SR-65, pp.33-62.

Tuller, B., Kelso, J.A.S. and Harris, K.S., (1981b). 'Phase Relationships Among Articulator Muscles as a Function of Speaking Rate and Stress,' Haskins Laboratories, Status Report on Speech Research, SR-65, pp.63-90.

Umeda, N., (1977). 'Consonant Duration in American English,' J. Acoust. Soc. Am., vol.61, pp.846-858.

Van den Berg, Jw., (1956). 'Direct and Indirect Determination of the Mean Subglottic Pressure,' Folia Phoniatrica, vol.8, pp.1-24.

Van den Berg, Jw., (1958). 'Myoelastic-Aerodynamic Theory of Voice Production,' J. Speech Hear Res., vol.1, pp.227-244.

Van den Berg, Jw., Zantema, J.T. and Doornenbal, P., (1957). 'On the Air Resistance and the Bernoulli Effect of the Human Larynx,' J. Acoust. Soc. Am., vol.29, pp.626-631.

Van Ess, J., (1938). The Spoken Arabic of Iraq. (2nd ed.). London : Oxford University Press. First edition appeared in 1917.

Warren, D.W., (1964). 'Velopharyngeal Orifice Size and Upper Pharyngeal Pressure - Flow Patterns in Normal Speech,' Plastic and Reconstructive Surgery, vol.33, pp.148-162.

Warren, D.W., (1976). 'Aerodynamics of Speech Productions,' in Contemporary Issues in Experimental Phonetics, edited by Lass, N.J.; New York : Academic Press, pp.105-137.

Warren, D.W. and DuBois, A.B., (1964). 'A Pressure-Flow Technique for Measuring Velopharyngeal Orifice Area During Continuous Speech,' Cleft Palate Journal, vol. 1, pp.52-71.

Warren, D.W. and Devereux, J.L., (1966). 'An Analog Study of Cleft Palate Speech,' Cleft Palate Journal, vol.3, pp.103-114.

Warren, D.W. and Wood, M.J., (1969). 'Respiratory Volumes in Normal Speech: A Possible Reason for Intraoral Pressure Differences Among Voiced and Voiceless Consonants,' J. Acoust. Soc. Am., vol. 45, pp.466-469.

Warren, D.W. and Hall, D.J., (1973). 'Glottal Activity and Intraoral Pressure During Stop Consonant Production,' Folia Phoniatrica, vol. 25, pp.121-129.

Wickens, G.M., (1980). Arabic Grammar : A First Workbook. Cambridge; Cambridge University Press.

Witting, C., (1953). 'New Techniques of Palatography,' Studia Linguistica, vol.7, pp.54-68.

Woodhead, D.R. and Beene, W., (1967). A Dictionary of Iraqi Arabic: Arabic-English. (eds.) Washington, D.C. : Georgetown University Press.

Wright, T.W., (1974). 'Temporal Interactions Within a Phrase and Sentence Context,' J. Acoust. Soc. Am., vol.56 pp.1258-1265.

Wright, W., (1962). A Grammar of the Arabic Language. (3rd ed.) Cambridge : Cambridge University Press.

Ziadeh, F.J. and Winder, R.B., (1957). An Introduction to Modern Arabic. Princeton : Princeton University Press.

Zimmerman, S.A. and Sapon, S.M., (1958). 'Note on Vowel Duration Seen Cross-Linguistically,' J. Acoust. Soc. Am., vol.30, pp.152-153.