PRODUCTION AND PERCEPTION OF LIBYAN ARABIC VOWELS

By

Albashir Ahmed

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy,
School of Education, Communication, and Language Sciences
Newcastle University

2008

Abstract

This study investigates the production and perception of Libyan Arabic (LA) vowels by native speakers and the relation between these major aspects of speech. The aim was to provide a detailed acoustic and auditory description of the vowels available in the LA inventory and to compare the phonetic features of these vowels with those of other Arabic varieties.

A review of the relevant literature showed that the LA dialect has not been investigated experimentally. The small number of studies conducted in the last few decades have been based mainly on impressionistic accounts. This study consists of two main investigations: one concerned with vowel production and the other with vowel perception. In terms of production, the study focused on gathering the data necessary to define the vowel inventory of the dialect and to explore the qualitative and quantitative characteristics of the vowels contained in this inventory. Twenty native speakers of LA were recorded while reading target monosyllabic words in carrier sentences. Acoustic and auditory analyses were used in order to provide a fairly comprehensive and objective description of the vocalic system of LA. The results showed that phonologically short and long Arabic vowels vary significantly in quality as well as quantity; a finding which is increasingly being reported in experimental studies of other Arabic dialects. Short vowels in LA tend to be more centralised than has been reported for other Arabic vowels, especially with regards to short /a/. The study also looked at the effect of voicing in

neighbouring consonants and vowel height on vowel duration, and the findings were compared to those of other varieties/languages.

The perception part of the study explored the extent to which listeners use the same acoustic cues of length and quality in vowel perception that are evident in their production. This involved the use of continua from synthesised vowels which varied along duration and/or formant frequency dimensions. The continua were randomised and played to 20 native listeners who took part in an identification task. The results show that, when it comes to perception, Arabic listeners still rely mainly on quantity for the distinction between phonologically long and short vowels. That is, when presented with stimuli containing conflicting acoustic cues (formant frequencies that are typical of long vowels but with short duration or formant frequencies that are typical of short vowels but with long duration), listeners reacted consistently to duration rather than formant frequency.

The results of both parts of the study provided some understanding of the LA vowel system. The production data allowed for a detailed description of the phonetic characteristics of LA vowels, and the acoustic space that they occupy was compared with those of other Arabic varieties. The perception data showed that production and perception do not always go hand in hand and that primary acoustic cues for the identification of vowels are dialect- and language-specific.

Declaration

I certify that all the material submitted in this work which is not my own work has been identified and that no material is included which has been submitted for any other award or qualification.

Signed:

Albashir Ahmed

Date: 15th April 2008.

Dedication

This work is dedicated to the memory of my late father, and to my beloved mother. It is also dedicated to my brothers, my sisters and my children.

Acknowledgements

The successful completion of this work is the result of the immense encouragement and help of many people. In particular, I appreciate the support given by my supervisor Dr Ghada Khattab; a very big "thank you" for your encouragement and patience during the preparation of this thesis. My heartfelt thanks are also due to the viva members Prof. Gerard Docherty and Dr Daniel Newman for their contribution in revising the entire thesis and their valuable comments. I am also indebted to Philip Harrison and Frantz Clermont who worked hard to prepare the synthesised stimuli for the perception task. I have appreciated the help of Dr Chris Letts, the Technical Site Manager and Safety Officer in the School of Education, Communication and Language Sciences, who gave me valuable advice regarding the best equipment needed for the production experiment and who also provided technical help during the work on the perception experiment. My thanks also go to my colleague Hussin Kriba with whom I shared ideas and thoughts related to the subjects we both studied and to the field of phonetics in general. Last but not least, I would like to thank all the participants in the experiments conducted in this study.

Table of contents

Tables

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

Figures

CHAPTER I

CHAPTER II

CHAPTER III

CHAPTER IV

CHAPTER V

CHAPTER VI

Acronyms and Abbreviations

1. Acronyms

AIT: Acoustic Invariance Theory

CA: Classical Arabic

DA: Derna Arabic

DRT: Direct Realist Theory

FLMP: Fuzzy Logical Model of Perception

LA: Libyan Arabic

MSA: Modern Standard Arabic

MT: Motor Theory

PT: Propagation Theory

QT: Quantal Theory

TA: Tripolitanian Arabic

SFT: Source-Filter Theory

2. Abbreviations

C: Consonant

Imp: Imperative

N: Noun

Pro: Pronoun

V: Vowel, Verb (depending on context)

CHAPTER I

INTRODUCTION

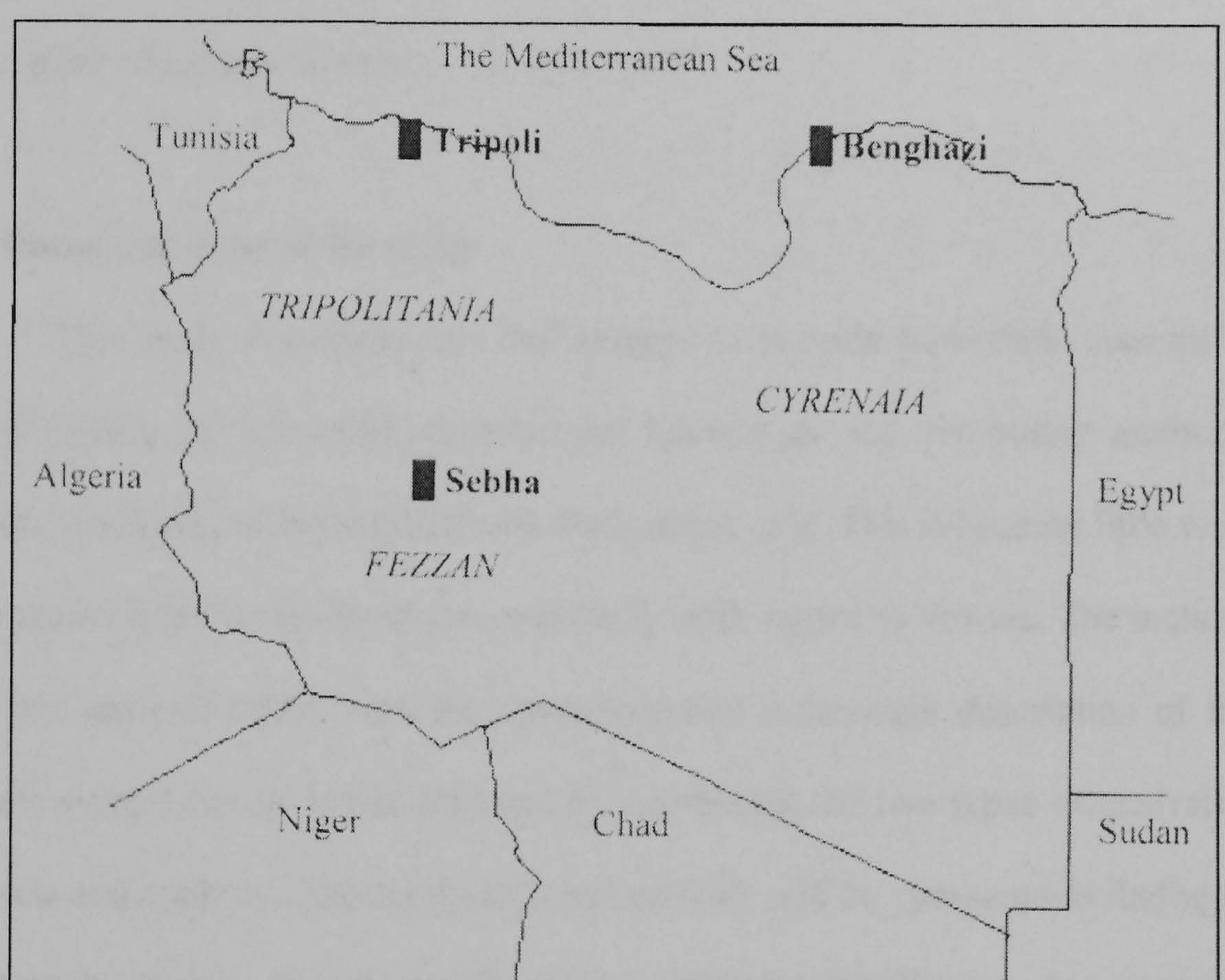
1.1. Area and topic

Arabic is one of the world's languages that has received considerable interest by researchers. Previous research has covered all aspects of the language including its syntax, semantics and phonology. Comparative studies of various aspects have also been conducted for different purposes. However, the comparative work has mainly been concerned with comparing the standard variety of Arabic with other languages. Standard Arabic is the formal variety alongside which a vast number of dialects coexist. These dialects are informally dominant and are used in everyday speech in a situation known as diglossia (Ferguson 1959), where two varieties of a language are used side by side for different but complementary purposes. Therefore, these Arabic dialects have also received some attention by researchers, though not to the extent standard Arabic has (for example, Alghamdi 1998, Al-Masri & Jongman 2004, Al-Tamimi, Carre & Marsico 2004).

Libyan Arabic (henceforth LA) is one such dialect that has been studied by several researchers, especially Libyan researchers (for example, Aurayieth 1982, Laradi 1983, Abumdas 1985, Elgadi 1986, Botagga 1991, and Al-Ageli 1995). However, most of the studies concerned are of the comparative type, comparing the dialect to other languages, mainly English, for pedagogical purposes. This involves outlining difficulties that Libyan learners face when learning English. Other studies are highly theoretical, dealing with the dialect using the generative theory framework (see, for example, Aurayieth 1982, Abumdas 1985, Elgadi 1986).

Moreover, the small number of experimental phonetic studies reported on Libyan Arabic are mainly based on auditory analyses of the dialect. Therefore, considerable disagreement is found among those researchers about some of the sounds used and their characteristics (see Chapter Three, sections 3.2 and 3.3). However, the fact that these differences might be partially related to the wide dialectal variation throughout the country should not be underestimated. Libya occupies a very vast area and therefore has a great number of dialects in spite of the small size of the population (less than 6 million).[1]

Fig.1. 1 A map of Libya showing the three major dialectal regions. [2]



---

[1] Libya 2003 census of population. http://en.ljbc.net/online/lypop.htm. [Accessed on 27.2.2008].

[2] The map was adapted from http://geography.about.com/library/blank/blxlibya.htm [Accessed on 12. 12. 2005].

There are three main dialect areas in Libya (Figure 1.1). These major dialects are the one spoken in Tripolitania including the capital city of Tripoli, the one spoken in Cyrenaia including Benghazi, the second major city in Libya, and finally the dialect dominating Fezzan around the city of Sebha in the south of the country.

To the best of the present researcher's knowledge, there have been no acoustic studies on any of these dialects. However, Laradi (1983) conducted an experimental study which focused on the physiological and articulatory aspects of the phenomena of pharyngealization in Tripolitanian Arabic (henceforth TA), the Arabic dialect spoken in Tripoli. The effect of pharyngealization on vowels was discussed by Laradi (see Chapter Three section 3.2), but no study has yet concentrated on a full acoustic description of Libyan vowels.

1.2. Focus and aims of the study

This study constitutes the first attempt to provide a phonetic account of the vowel system of LA using experimental techniques and combining auditory and acoustic methods, with more focus on the acoustic side. This is because little work has been done on the acoustics of LA, especially with regard to vowels. The inclusion of auditory analysis arises from the contention that a thorough description of the LA vocalic system can be better achieved by combining the two types of analysis: both acoustic and auditory. The results of these analyses will be compared to findings from other Arabic dialects in order to discover similarities and differences in vowel patterns between various Arabic dialects.

1.3. Importance of the study

This study provides a comprehensive account of the vocalic system of LA and outlines the phonetics of and some of the phonological differences among the vowels contained in this system.

Another contribution made by this study is its use of relatively objective tools of data collection and data analysis. These tools include the use of computer-based analyses which enable the researcher to avoid relying mainly on the auditory examination of material produced by participants. According to Docherty and Foulkes (2000, p113) such auditory judgement of speech may lead to inaccurate transcription of data for analysis.

Another advantage of this study is the combination of speech production and perception. Integrating these two aspects of human speech prevents researchers from only replying on participants' output when drawing conclusions about vowel properties, as is the case in most previous related studies. In this respect, listeners' reactions to acoustic cues that are characteristic of particular vowels are investigated to find out to what extent vowel quality and quantity are important in the identification of these vowels.

The results obtained from this study may be used by researchers working on the LA dialect for different purposes, such as cross-dialectal and cross-linguistic comparisons involving vowel inventory size and realisations of the vowels contained in the inventories of the dialects and languages under investigation. The results may also be used by practitioners for language teaching purposes. The teaching of Arabic dialects (as opposed to the Standard) is becoming more and more popular, and it is important for language teachers to be aware of the vowel quality and quantity aspects of LA and to what extent these vowels may have more in common with languages like English than the Standard Arabic variety.

1.4. Research questions

As mentioned above, this research investigates the production and the perception of LA vowels by native speakers of the dialect. With regard to production, the focus is on the number of vowels contained in the dialect inventory and the qualitative and quantitative differences between them. As for perception, the research questions are oriented towards the reactions of native listeners of the dialect to differences that might be found which are relevant to vowel production and whether or not these differences are language-specific. The main research question is stated below followed by four more specific questions.

Main question

What is the vowel inventory of Libyan Arabic and how are the vowels contained in this inventory produced and perceived by native speakers of the dialect?

Specific questions

1. What is the phonological inventory of vowels available in LA?
2. How are these vowels realised qualitatively and quantitatively in this dialect?
3. Which acoustic cues do native speakers of LA respond to in perceiving vowels?
4. To what extent does attention to varying acoustic cues in the perception of these vowels relate to the acoustic cues that are found relevant in production?

1.5. Methods

The answer to the specific research questions above requires the adoption of a multi-method approach to achieve the study's aims. The approach used encompasses a literature review, native speaker intuition and experimental investigation. A detailed account of the methods employed is given in Chapters Four and Five. However, a brief overview of these methods is given below:

1. What is the phonological inventory of vowels available in LA?

The answer to this question requires investigation of the relevant literature to determine the number of vowels available in LA and their realisations. The researcher will also rely on his native speaker intuition where information from the literature is lacking.

2. How are these vowels realised qualitatively and quantitatively in this dialect?

Once the number of vowels available in the dialect is obtained, a representative sample consisting of target words containing these vowels will be designed in order to be elicited from a sample of native speakers of the dialect. After recording the production of the participants, the data will be analysed acoustically and auditorily to determine the features these vowels have and their variations across and within speakers.

It has been observed by some researchers that Arabic short vowels do not only differ from their long counterparts in quantity but also in quality (Alghamdi 1998, Ghazeli 1977). However, another view maintains that this difference might not apply to the same degree in Modern Standard Arabic (henceforth MSA). In fact some researchers claim that short vowels and long counterparts only differ in quantity and not in quality in MSA (see, for example, Muftah, 2001, p.79). Therefore, an investigation into this matter will be conducted in terms of production to determine how the quality and quantity of vowels interact with each other in LA vowels.

3. Which acoustic cues do native speakers of LA respond to in perceiving vowels?

Once the acoustic patterns of LA vowels produced by native speakers are specified, a set of perception experiments involving synthesised vowels will be conducted in order to find out what role different acoustic cues play in native listeners' perception. Specifically, the acoustic cue weighting of vowel quality and quantity will be investigated in order to find out how these cues interact in perception and which is more important for LA listeners.

4. To what extent does attention to varying acoustic cues in the perception of these vowels relate to the acoustic cues that are found relevant in production?

This question aims at exploring the relationship between the production and perception of LA vowels. One aspect of this relationship that has not been sufficiently studied by researchers into Arabic and its dialects is how the production and perception of the same vowel interact in a speaker (see, for example, Al-Tamimi and Barkat-Defradas 2002; Barkat-Defradas et al 2003). Therefore, this relationship is investigated in order to find out whether listeners use the same acoustic cues found in production to perceive vowels and whether these cues are language-specific. In this respect, the production of some of the vowels produced by native speakers will be compared to their perception.

1.6. Organization of the study

The thesis proceeds from a review of the relevant literature to the methodology used and the analysis of the results, ending with a discussion of the results in the light of the literature reviewed. Accordingly, apart from this introduction chapter, the thesis comprises five other chapters.

Chapter Two is the first part of the literature review and is devoted to dealing with the production and perception of speech. Since these two aspects constitute the focus of the study, it is necessary to clarify their nature and to see to what extent they are related. Some of the concepts that are thought to be relevant to the study are explored, such as vocalic contrast, coarticulation, categorical perception and normalisation. Problems of speech perception such as lack of invariance and linearity are also discussed. Theories of speech production and speech perception are also reviewed in this chapter in order to obtain a suitable theoretical framework within which this study can be conducted. The aim is to develop the theoretical background necessary for discussing and explaining the results obtained from the data elicited from participants. This eventually helps also in finding reasonable interpretations of the findings of the study.

Chapter Three is devoted to reviewing literature related to the Arabic language and particularly LA. Although, LA is the main concern of this work, it was inevitable that other Arabic varieties needed to be discussed, especially the standard variety of Arabic, namely MSA. The reason behind this is the fact that these varieties are interrelated, and comparing them helps to specify their similarities and differences especially with regard to vowel inventory size and the phonetic features of the vowels contained in the inventories of these dialects and varieties. The review of literature related to LA and other Arabic varieties and dialects also helps to identify gaps in the research which need to be filled; and to build on previous experience, especially that related to research methodology.

The beginning sections of Chapters Four and Five present the methods used to obtain the data from native participants of the dialect. Different types of data require different methods. Therefore, two types of elicitation techniques are used. Production and perception data are each elicited from twenty native speakers. Efforts are made to avoid the limitations of previous studies as much as possible by, for example, using technically sophisticated equipment. Also, care is taken in choosing the sample and in the material and procedure used in order to be appropriate for each task and to lead to reasonably reliable and valid findings.

The second halves of Chapters Four and Five present and discuss the results obtained from the production and perception studies. In addition to the analysis and discussion of vowel formants and duration, factors such as inter- and intra-speaker variability, vowel quality overlap and the interaction between vowel quality and quantity are considered in terms of both production and perception. In the discussion of the findings related to these topics, reference is made to findings from studies on other Arabic varieties in order to find out how LA vowels compare with those of other dialects.

Chapter Six is devoted to a general discussion and conclusions. The aim of the general discussion is to discuss the research questions in the light of the results obtained. This enables us to assess the answers obtained to these questions and to review the aims and achievements of the study. In the light of this review and assessment, the theoretical and practical implications of the findings are then considered. Finally, the limitations of the study are reviewed and recommendations for future research are made.

PART ONE

REVIEW OF LITERATURE

CHAPTER II
SPEECH PRODUCTION AND PERCEPTION

CHAPTER III
ARABIC VOCALIC SYSTEM

CHAPTER II

SPEECH PRODUCTION AND PERCEPTION

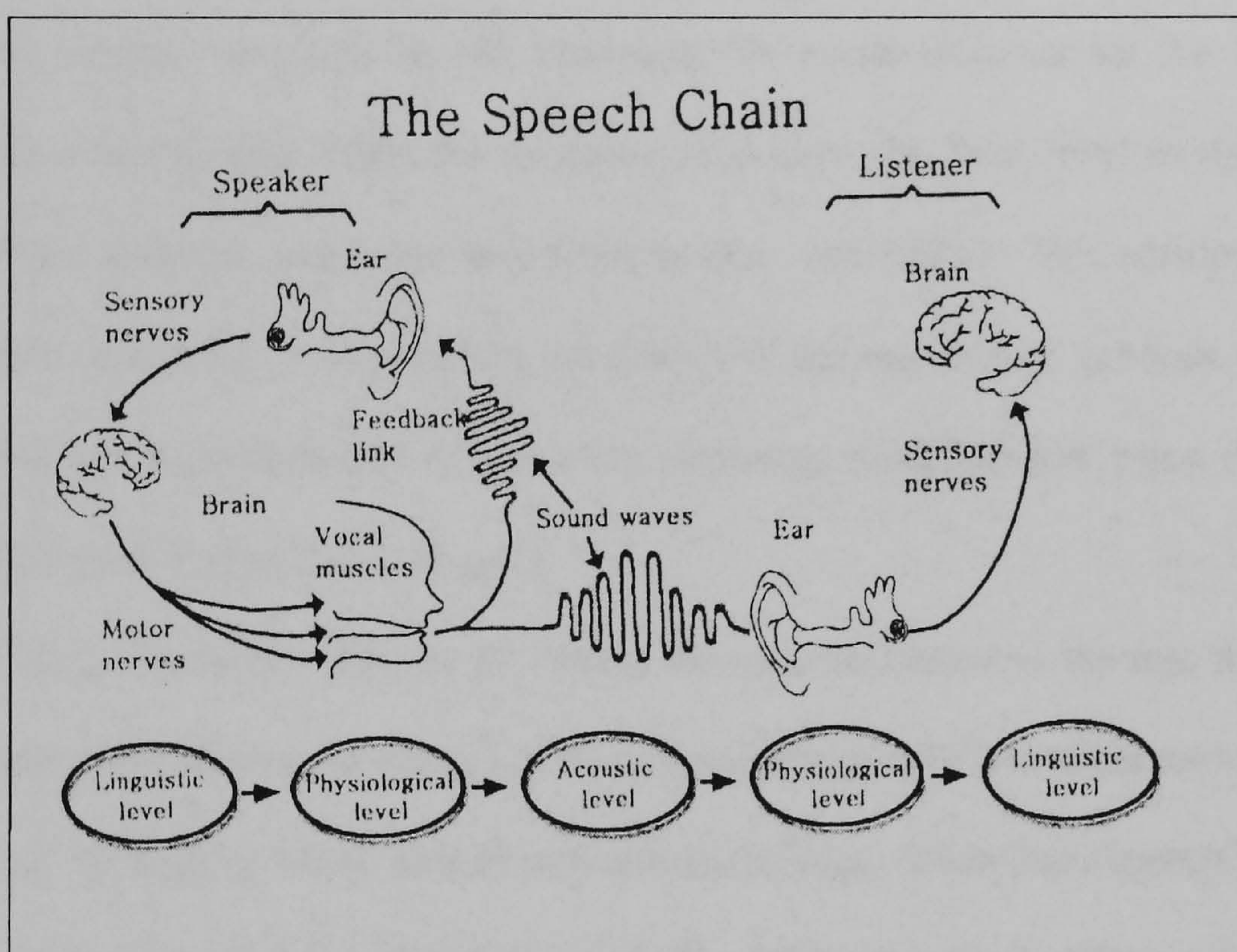
2.0. Introduction

Production and perception constitute two important aspects of speech and are closely related. From a simplistic point of view, the principal aim of a speaker when they produce an utterance is to be perceived by a listener (Tatham & Morton, 2006, p. 218) and the sound contrasts produced by a speaker should in principle correspond to ones perceived by the listener (Stevens 1972). After all, listeners are to perceive the language forms that are produced by speakers (Fowler & Galantucci 2005, p.634) in order to interpret the intended message. However, speech is highly variable due to factors such as speech rate, speaker age, physiology, gender, dialect, social class, and emotional state. As a result, listeners do not only rely on the acoustic and auditory signal in order to understand speech, but also make use of other visual, indexical, and contextual clues in order to process the speech signal. Theories of speech perception have dealt with the problem of the richness of the signal in various ways and have suggested different units of encoding of the speech signal and different types of phonological abstraction. Within vowel perception work, one debate has revolved around whether vocalic targets are static or dynamic, and on which types of information provide better cues for vowel identification and categorisation (Strange 1989). In the following sections of this chapter, the literature related to speech production and perception is reviewed.

2.1. The nature of speech production and perception

The ability to produce and perceive speech is taken for granted. However, the processes involved are highly complex and entail a number of actions from both the speaker and the listener. In fact, speech passes through a chain consisting of five levels known as the speech chain (Denes & Pinson 1993, p.1; see Figure 2.1).

Fig.2. 1 The speech chain: the stages of spoken communication (Denes & Pinson 1993, p.1)



This speech chain starts from the brain of the speaker and ends at the brain of the listener. At the first level, which is linguistic, the message intended by the speaker is arranged in a linguistic form, i.e. in words and sentences. Then the brain sends commands through the nerves to the muscles of the vocal organs which become activated and produce the sounds and utterances. This process represents the second level in the speech chain which is physiological in nature. These movements of the vocal organs when utterances are produced cause a change in the air pressure which generates a speech sound wave that moves through the air from the speaker to the perceiver. This third level in the speech chain is physical and is specifically called the acoustic level. The hearing system, including the ear and the nerves transmitting the signal to the brain, represents the fourth level in the speech chain which is physiological in nature. The ear is sensitive to the change in the air pressure, and the hearing mechanism is stimulated. As a result, nerve impulses are produced and which reach the human brain where a lot of complex nerve activities occur in order for the speaker's message to be perceived. Decoding the nerve impulses by the listener's brain into a meaningful linguistic utterance represents the final level in the speech chain. When speakers speak they also listen to their own speech. This enables them to assess their production by comparing the quality of the sounds they produce with that they intended to produce and to make the necessary modifications when things go wrong (Denes & Pinson 1993, pp.1-4).

One of the major reasons for speech being so important is the role it plays in the development of human culture. Through speech, people can communicate with each other by sharing ideas, experience and knowledge. Therefore, speech deserves careful study because it is the major means by which humans communicate and by investigating this means, human communication can be developed (Denes & Pinson 1993, p.2).

2.2. Variability in speech production

A substantial amount of research in speech production has shown that speech signals vary considerably. Variation in speech production might be explained by several factors. These include within-speaker and between-speakers variation (Perkell 1990, p.266). The latter includes idiosyncratic and social-indexical differences as well as differences in vocal tract size as a result of age and sex differences. Therefore, two different speakers may articulate the same intended sound with different acoustic patterns and different sounds with the same acoustic patterns (Peterson and Barney 1952, Fant 1973).

Intra-speaker variation may result from contextual factors such as speaking rate and phonetic context. When vowels are contiguous to consonants in normal speech these vowels are affected by the properties of these consonants. This is known as coarticulation, another consequence of which is that vowels are "undershot" (Strange 1989, p. 2082). In other words, the target vowel is never reached when it is produced in continuous speech due to factors such as phonetic context, stress and speaking rate. One consequence of this is that the "vowel space populated by tokens of a particular vowel type produced in ... [different] contexts will often overlap significantly with regions containing tokens of other vowel types" (Strange et al 1983 p.695). In spite of this articulatory and acoustic variability, listeners are still able to perceive vowels as intended by speakers. It is argued that listeners compensate for this variability in order to perceive the intended vowel targets through a process called normalisation (see section 2.7.3).

Coarticulation is seen by some researchers (e.g. Liberman et al 1967) as a fundamental characteristic of speech, and one by which information can be conveyed rapidly. It refers to the effect of sounds on neighbouring sounds when they are grouped together in an utterance. Therefore, coarticulation implies that some features of a phonetic sound are found within an adjacent sound. In fact, coarticulation has a major impact on the characteristics of both vowels and consonants which is also reflected in the characteristics of the syllables they constitute (Massaro 1992, p.82). In a study conducted by Lindblom (1963), for example, it was found that when syllables are produced quickly, their duration becomes shorter. This in turn leads to more coarticulation between vowels and consonants and weaker segment boundaries. However, it is still possible to detect areas of transitions from one segment to another. Major spectral changes in the speech signal can still be located using acoustic analysis procedures which make it possible to recover the articulated segments from the acoustic signal (Lofqvist 1990, p. 291).

To sum up, speech variation is found within and across speakers due to several factors including speaker, gender, age, dialect and social class in addition to phonetic context. Therefore, when dealing with speech data it is crucial to control these different sources of variability and to consider the possible effects of variation arising from factors that are hard to control (Perkell 1990, p. 266).

2.3. Variability in vowel production

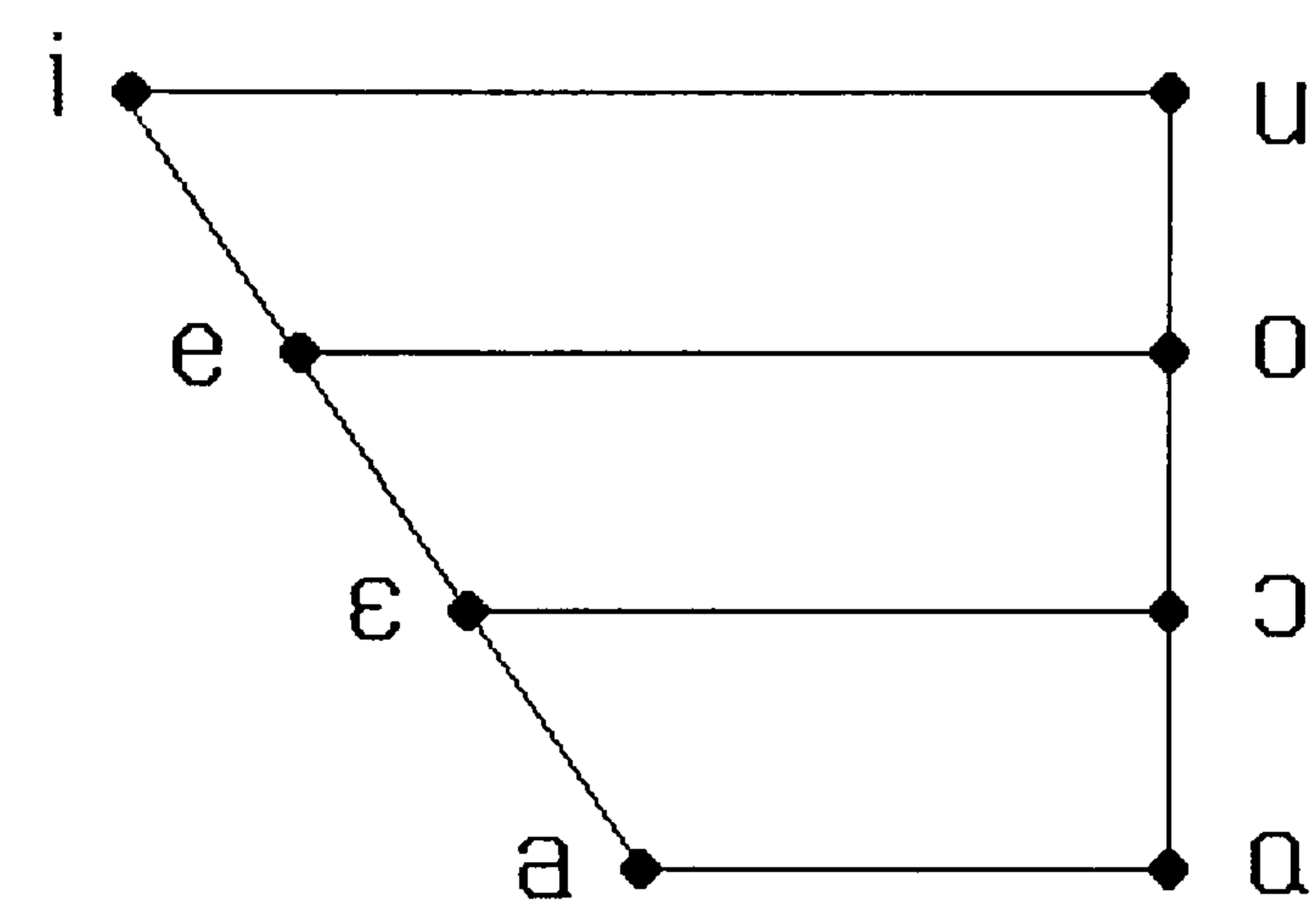
Vowel inventories of the world's languages show considerable variability in terms of size and distribution in the acoustic vowel space. According to Maddieson (1984), the vowel inventory size in the languages of the world ranges from 3 to 24 distinct vowels which also vary in their features. However, research has found that the vowel systems of these languages share universal tendencies. Studying these tendencies is considered crucial for any theoretical perspective in phonetics (Schwartz, Boe, Vallee and Abry 1997). For instance, the most common vowel system consists of 5 vowels. Moreover, certain features and feature combinations tend to be more common than others (Ladefoged & Maddieson 1996). For example, the vowel qualities /i, e, a, o, u/ are more common than others and are therefore considered a model vowel inventory in the languages of the world. An explanation for this was offered by Lang and Ohala (1996) who investigated English vowels and found that these vowels have auditorily robust features. Robust features of vowels are those which can be identified in a relatively shorter time window (50ms or less) compared to non-robust features which need a longer time to perceive (Carrie &Ohala 1996, p. 4). These robust features found in these five vowels make them universally more common. That is, they are more frequent in the languages of the world than other vowels.

As for the distribution of vowels in the acoustic vowel space, Al Tamimi and Ferragne (2005) investigated the effect of vowel inventory size on vowel space size in three languages which differ in their vowel systems. These languages are French, Moroccan Arabic and Jordanian Arabic. They found that what they call point vowels – /i, a, u/ – are approximately in the same positions in all three languages when these vowels are produced in isolation and in syllables, but not when in words. This was explained by the H&H (hyper- and hypo-speech) hypothesis (Lindblom 1989), which maintains that speakers rely more on context, and thus give less information when producing vowels in words which leads to vowel undershoot phenomena. However, the acoustic vowel space was found to correlate positively with inventory size. That is, the larger the inventory, the larger the vowel space. Therefore, the acoustic vowel space was found to be larger in French, which has 11 vowels, than in Jordanian and Moroccan Arabic which have 8 and 5 vowels respectively. The Jordanian acoustic vowel space was also found to be larger than the Moroccan acoustic vowels space.

2.4. Vowel classification and theories of vowel production

Vowel inventories are traditionally represented in terms of the cardinal values on a quadrilateral, first suggested by Daniel Jones (1964) to work as a standard reference system for transcribing and classifying vowels. The quadrilateral represents the area of the mouth in which the tongue can move up and down, and forward and backward to produce all speech vowels, and beyond which no vowel sound can be produced.

Fig.2. 2 Primary Cardinal Vowels (Jones 1964)



As seen in Figure 2.2, this quadrilateral has two dimensions: vertical and horizontal. The vertical dimension represents tongue height while the horizontal dimension represents tongue advancement during the production of vowels. Therefore, the vowel /i/ is considered a high front vowel, the vowel /u/ is a high back vowel and the vowel /a/ is described as a low front vowel. In addition to tongue height and tongue advancement, vowels are also characterised by the two lip positions (Ladefoged 2001, p. 40). For example, the high front cardinal vowel /i/ is produced with unrounded lips while these take a rounded position in the articulation of the high back cardinal vowel /u/. Besides this articulatory classification, vowels can also be described acoustically, that is, in terms of formant frequencies resulting from resonance in the oral cavity during the production of these vowels. This resonance and the formant frequencies are captured in what is known as the Source-Filter Theory, the focus of the following section.

2.4.1 The Source-Filter Theory

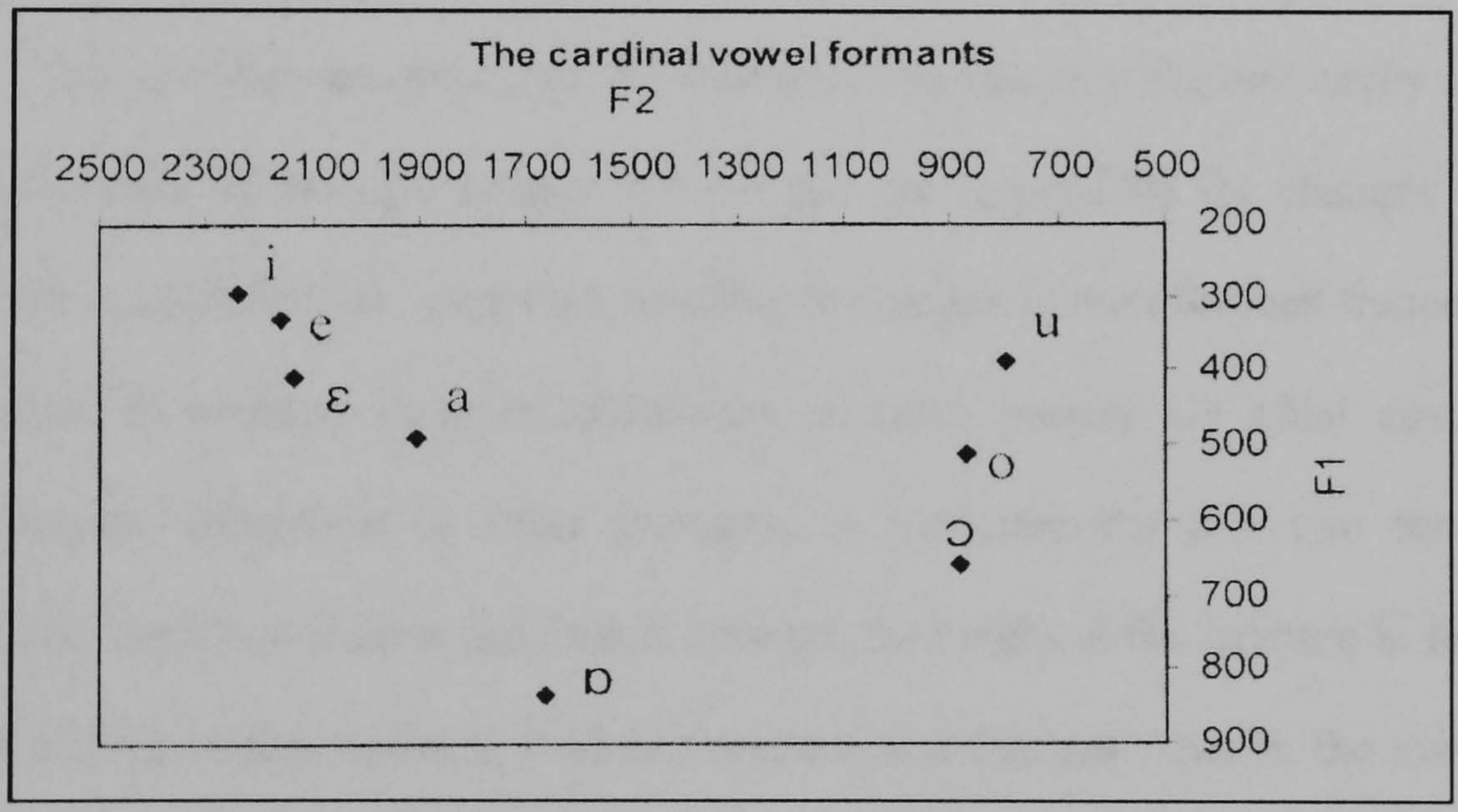
The Source-Filter Theory (henceforth SFT) was proposed by Fant (1960) in order to explain the production of vowels. According to this theory, two factors are involved in the production of vowels: source and filter. The vocal folds are considered the source of the voice produced and determine the amount of air that passes from the lungs up to the oral cavity through the larynx. The oral cavity is the second factor in the production of vocalic sounds. Adjustments in the shape of this oral cavity result in different types of resonance that produce different vowels.

Although sounds can be analysed in terms of these two factors, it is the filter which determines the type of vowel produced. Within a speaker, the source is consistent in all the vowels produced due to the relatively constant state of the source or the vocal folds; all vowels are produced with vocal fold vibration. However, the filter can be shaped to produce different types of vowels. It is changes in the position of the tongue, lips and jaw that lead to the production of a variety of distinctive vocalic sounds that have different frequency values (Rosner & Pickering, 1994, p.3).

In this respect the vocal tract is divided into two resonating chambers which are directly responsible for the production of the first two formants. Resonance in the back of the vocal tract, which is from the larynx to the top of the throat, produces the first formant frequency; and resonance in the front of the vocal tract, which is from the top of the throat to the lips, generates the second formant frequency. By analysing these frequencies, vowels can be described and identified. According to Peterson (1951), the first and second formants matter most in determining the quality of vowels. The frequencies of these first two formants form the acoustic cues that are often adequate for perceptually recognising a certain vowel (Strange 1989, p. 2081).

Subsequently, the quadrilateral (see section 2.3) that is traditionally used to represent vowels was later adapted by Joos (1948) to represent vowels acoustically in a two dimensional plane which represents the first two vowel formants, namely F1 and F2. In this plane, F1 and F2 are represented by the vertical and the horizontal axes respectively. These two axes meet at the right hand upper corner in order for the plane to resemble the traditional quadrilateral, which reflects the oral cavity where vowel articulations take place and where the two formant frequencies are produced (Rosner & Pickering, 1994, p.11). The configuration of the acoustic plane was first suggested in order for it to show the parallel link between vowel acoustics and vowel production as shown in Figure 2.3 below.

Fig.2. 3 Cardinal vowel formants as produced by Daniel Jones[3]



[3] The production of the cardinal vowels was taken from Ladefoged's Course in Phonetics, found online at http://www.phonetics.ucla.edu/course/chapter9/cardinal/cardinal.html, retrieved on 21.9.2007. Then formants for cardinal vowels were measured and plotted as shown in figure 2.3.

Although the F1/F2 plane is supposed to reflect the acoustic representations of tongue movement in the oral cavity, these representations are also affected by other articulatory factors such as the protrusion of the lips. The latter lowers formants as a result of making the vocal tract size bigger. In addition, the plane is not capable of showing vowel duration, which is a distinctive feature in the vowel systems of some languages. Finally, vowel representation in the F1/F2 plane is also affected by differences among speakers as a result of variations in vocal tract length and dialectal, social and generational differences (Rosner & Pickering 1994, p.15). Therefore, the mismatch between the cardinal vowels represented by the quadrilateral and the same vowels represented by a formant plot as produced by Daniel Jones himself is obvious, especially with regard to the cardinal vowel /ɒ/ which is supposed to be further back to the right side of the plane.

2.4.2 Vocalic contrasts

When vowels are produced, the changes in the shape of the oral cavity caused by movements of the lips, tongue and the jaw are responsible for changes in the resonance qualities of the vocal tract resulting in changes in their formant frequencies. Therefore, in addition to their articulatory contrast, vowels are often contrasted according to differences in these formants, in particular the first two formants. Generally, there is a reverse connection between the height of the stricture in the oral cavity when a certain vowel is produced and the first formant. That is, the lower the stricture in the mouth the higher the first formant and vice versa. On the other hand, the second formant is often connected with the back-front dimension, i.e. the more to the front the stricture in the mouth, the higher the second formant. The second formant is also affected by the position of the lips. Lip rounding and protrusion increases the vocal tract length, thus lowering the second formant (Ladefoged 2001, p. 39-43). This relation between articulatory configurations and acoustic signal is captured by the Quantal Theory which is reviewed in the following section.

2.4.3 The Quantal Theory

The Quantal Theory (henceforth QT), which was proposed by Stevens (1972), maintains that the sound system of a language is determined by sound features and their combinations. These features also play a major role in the perception of these sounds. These features had been used widely in phonological descriptions of sound inventories (Clements 2003). For example, vowels were described in terms of their distinctive features as being ± high and ± back as far as their place of articulation in the oral cavity is concerned. However, in spite of the importance of phonological features in defining the structure of sound systems and in the perception of the sounds contained in these systems, their phonetic basis before QT remained less understood. QT attempted to integrate phonetics with phonology as far as sound features are concerned and sought to explain the relationship between articulatory configurations and acoustic output. It also aimed to explore why some sounds and sound systems are more common than others across languages. Each of these points is discussed further below.

QT gives equal importance to the articulatory and acoustic domains of speech. It maintains that the relationship between the articulatory and the acoustic domain is quantal. That is, when the vocal tract configuration changes, the acoustic output also changes accordingly. However, the acoustic parameter is more sensitive to some changes in the articulatory parameter than to other changes in the same articulatory parameter. That is, this relation between the two parameters is not linear. To use

Stevens' words, "the articulatory-acoustic relations are quantal in the sense that the acoustic pattern shows a change from one state to another as the articulatory parameter is varied through a range of values" (Stevens 1989, p. 357). For instance, the acoustic parameter shows a considerable change because of a small change in the articulatory configuration, while in other cases the acoustic parameter remains relatively stable when a similar change in the articulatory parameter occurs.

This non-linear articulatory-acoustic relationship is clearly illustrated in Figure 2.1 which represents a rapid shift in the acoustic parameter from one stable area to another for a linear articulatory movement. As stated before, sound features in the QT play a major role. An example provided by Sonderegger (2004, p. 14) shows that moving the tongue from the sound /s/ (region 1 in Figure 2.4) which has a [+anterior] feature to the sound /ʃ/ (region 3) which has a [-anterior] feature leads to a sudden transition between the two (region 2).

Fig.2. 4 Quantal relations between articulatory and acoustic parameters (Stevens 1989, p. 358)

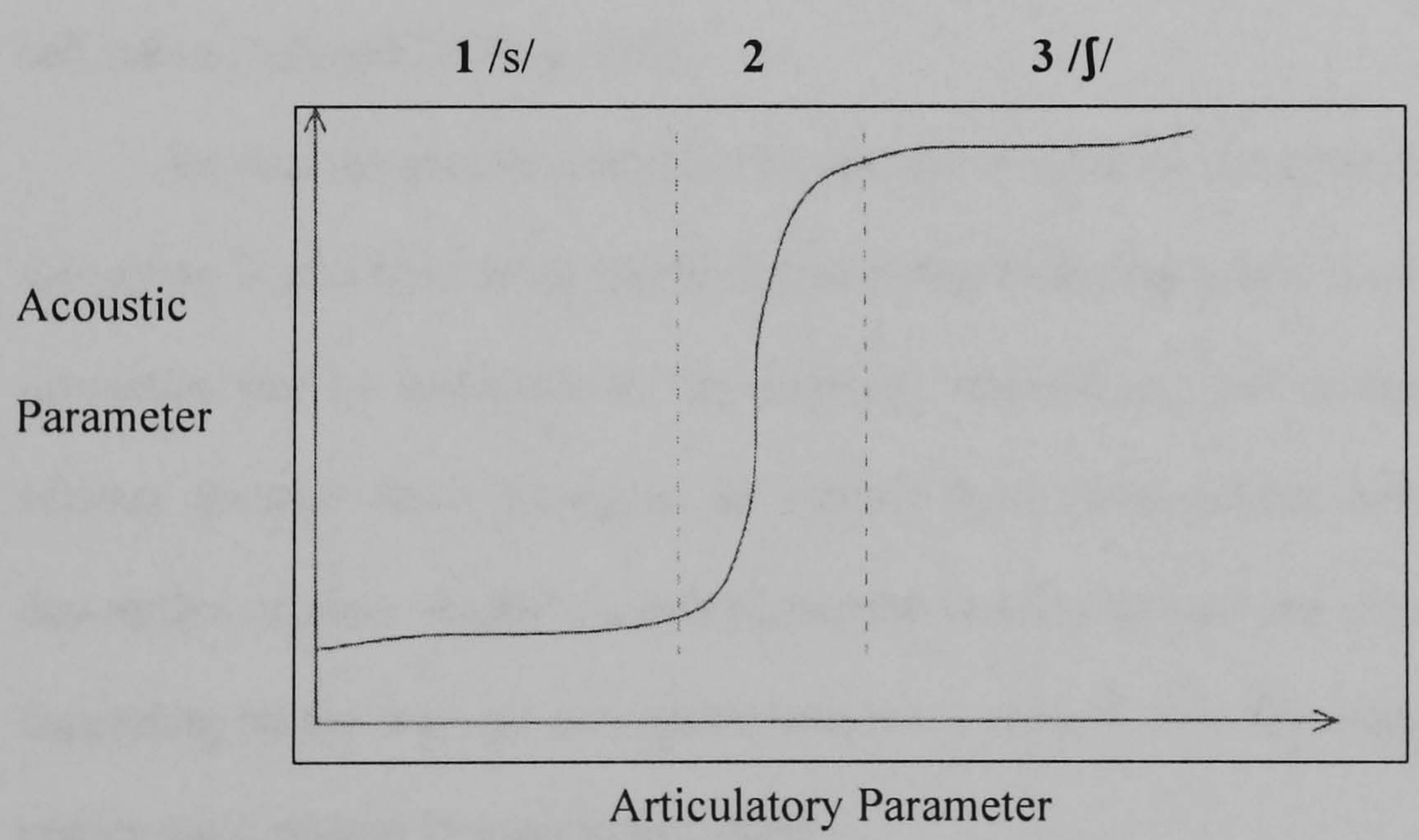


Figure 2.4 shows that the acoustic parameter is not sensitive to changes within an articulatory region while this parameter is more sensitive between regions of the

articulatory parameter. According to the QT, linguistic contrast involves differences between quantal regions, not within them. Therefore, continuous movement within a single quantal region will result in an acoustic steady state. Quantal regions are correlates of distinctive features which may include manner and place of articulation for vowels, for example, and preferred sound categories are those that occupy stable regions which are separated by unstable regions.

QT maintains that by relating oral vowel articulations to acoustic output, cross-linguistic trends in vowel inventories can be predicted (Johnson 2003). For example, the vowels /i, a, u/ are the most frequent vowels across languages (Maddieson 1984) because they are produced at regions of acoustic stability. The vowel /i/ is produced at the intersection of F1 and F2 created by a stricture in the palatal region and where the F2 value is stable. Vowel /a/ is articulated at the intersection of F1 and F2 where the front and back cavities have approximately the same lengths. Finally, vowel /u/ is produced at a region of stability for F1 near the soft palate (Johnson 2003, p. 112).

To sum up, speech is highly variable due to internal and external factors. This variability is exhibited at all levels of production including vowel production. Vowel properties can be examined at the auditory, articulatory, and acoustic levels and various theories have attempted to explain how these levels contribute to the description of these vowels. Speech perception is also captured in a variety of theories depending on the way the perception process is viewed. The following sections will review work related to speech perception.

2.5. Speech perception

Speech perception is a process that involves a communicative act in which a listener derives meaning from a speaker. It is a process by which a speaker of a language is understood by another speaker of that language. Deriving the meaning from speech may seem a straightforward process. However, many events take place at the physiological, acoustic and perceptual levels which make this process a lot more complex. According to one view of perception, the listeners' task may be seen as identifying the properties of the acoustically invariant signal which they map onto meaningful entries in the brain. This process would involve segmenting the signal into abstract linguistic units including phonemes and their features (Studdert-Kennedy 1976). According to this view, only minimal information that is crucial to the meaning is identified by the listener. These pieces of information are termed as speech cues. Additional linguistic information that is not relevant to the meaning is discarded. However, the mechanics of how that process occurs is still not clear.

For this perception process to be successful, the signal must be invariant and linear. Invariance has long been considered a condition for the speech signal to be perceived by listeners (Chomsky & Miller 1963). In practice, however, this condition is never met due to variability in the speech signal. As mentioned in section 2.2, the same utterance may vary due to several factors. These include inter- and intra-speaker variability due to vocal tract size, speaking rate, dialect, socio-economic factors, emotional state and the social context, among others (Pisoni & Lively 1995, p. 437). For the speech signal to be linear, all instances of the signal must have the same information and the linguistic units of that signal should be in a linear relation with each other with no overlap (Wright et al, 1997, p2). However, in reality, this condition is not met either. When speech is produced naturally the organs of speech move continuously and rapidly. Therefore, the relation among speech sounds when they are put together in speech streams is not linear. That is, sounds are no longer clearly defined and acoustically separate. Instead, they overlap and blend as a result of coarticulation. Therefore, their features will be affected and coloured by each other due to that overlapping and coarticulation. Accordingly, recovering these sounds from the speech stream will not be an easy task (Hockett 1955, p.55), and the information contained in any linguistic unit in the signal will show considerable change and overlap with adjacent units (Delattre et al 1955).

To solve these two major problems, speech perception has been viewed in different ways based on what listeners perceive and how they perceive it. Subsequently, a number of theories of perception have been constructed. The majority of these models aim at segmenting the speech signal into its components (Wright et al 1997, p.26) in an attempt to understand how the listener recovers the spoken and coarticulated sequence of phonological elements. Most theories suggest that there is a level of invariance in the signal that should be linked to some invariant symbol at the end of the perception process. However, decades of research on the relationship between invariance in the signal and invariance in the symbol have shown that this relation is not a simple one mapping phonetic features to acoustic signals (Strange et al 1983, p.695). This is partially caused by the inflexible nature of the phonetic symbols in representing speech sounds. Therefore, this relationship has been viewed differently in different speech models. While this relationship is direct in, for example, Direct Realism Theory; it is indirect in, for example, Motor Theory (Tatham & Moton 2006, p.213).

Theories of speech perception are classified into two major categories: strong auditory theories and strong articulatory theories (Neary 1997, p.3241). Strong auditory theories take the view that acoustic and auditory events are capable of explaining speech perception, while strong articulatory theories maintain that speech perception can be explained by articulation alone. The major model of the former category is Acoustic Invariance Theory, and the major ones of the latter category are Motor Theory and Direct-Realist Theory. Apart from these strong theories there are other weak approaches that have attempted to achieve a compromise between the two strong types of speech perception theories, such as the Fuzzy Logical Model which aims at integrating the information necessary for perception from different sources including auditory and visual input.

2.5.1. Acoustic Invariance Theory

Acoustic Invariance Theory (henceforth AIT) developed out of the Quantal Theory, as Stevens (1972) attempted to identify acoustic correlates for the distinctive features which were originally considered the basic units of perception and defined in articulatory terms. AIT aims at solving the variability problem by looking for invariant cues in the acoustic signal. The assumption is that in spite of the variability in speech, invariant cues do exist in the acoustic signal (Blumstein & Stevens 1980). According to this theory, distinctive articulatory features correspond to invariant properties in the acoustic signal. For consonants, this correspondence between articulatory and acoustic domains is manifest in the spectral changes that take place across segment boundaries. On the other hand, distinctive features for vowels are manifest in the acoustic signal in their steady state where the two formants remain relatively stable for a period of time (Hawkins 2004, p. 14). Therefore, perception is totally determined by the auditory processing of the acoustic signal irrespective of how this signal was produced by the speaker (Rosner & Pickering 1994, p.373). For example, the same vowel may have a variety of articulatory targets for different speakers. Yet, the corresponding acoustic signal still approximately bears invariant properties. This, for the proponents of this theory, this an indication that the linguistic properties that are relevant to vowel distinction are acoustic or auditory rather than articulatory (Nearey 1997, p.3242).

In order to address the issues of linearity and invariance, two theoretical concepts have developed out of this and other auditory-based theories. These are categorical perception and normalisation.

2.5.1.1. Categorical perception

Categorical perception suggests that sounds are grouped into different categories within the acoustic space. A listener is able to discriminate between sounds belonging to different categories and will perceive these sounds as one unit when they belong to a single category even when these sounds have different realisations (e.g., Liberman et al 1957, Ambramson & Lisker 1967, Pisoni 1971).

Liberman et al (1957) conducted an experiment in which a continuum of [b]-[d]-[g] was used. Listeners were able to make clear boundaries between these three sounds in spite of the gradual change of the F2 transition of the following vowel which was [eɪ]. Therefore, listeners reported hearing the words *bay, day* and *gay* in all of which the first sound represents a distinct sound category. However, those listeners were not able to discriminate instances from the same phonetic category.

Abramson and Lisker (1967) conducted a similar experiment using a /b/-/p/ continuum. Both consonants are bilabials and the only difference between the two is their voice onset time. /b/ has a shorter VOT than /p/. Participants in the experiment had no difficulty in identifying the stimuli. More interestingly, they did not report hearing something like /b/ or /p/ at the middle of the continua. Rather they heard either /b/ or /p/ and made a clear boundary between the two.

The idea of categorical perception can be further explained using Figure 2.5 below which represents eight stimuli that have equal distance apart along a certain dimension.

Fig.2. 5 Categorical perception[4]

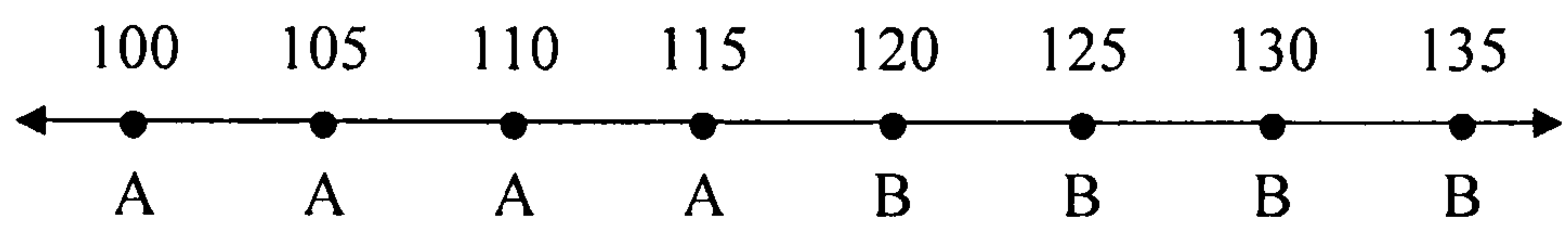


Figure 2.5 assumes that these points in the phonetic space correspond to two phonological categories. For instance, the stimuli from 100-115 units are perceived by listeners as category A while those from 120-135 are recognised as belonging to category B. The units may refer to any linguistic cues such as durational or spectral cues including formant and formant transitions. In this example, although the physical change between each pair of stimuli is the same, the perceptual distance between the two stimuli that have 115 and 120 unit values is different. That is, each one of these two stimuli is recognised as belonging to a different category.

The idea of categorical perception is in harmony with the concept of the 'perceptual magnet' introduced by Kuhl (1991). The perceptual magnet refers to a certain location in the vowel space around which prototypes or the best instances (Rosch 1975) of the vowel as perceived by listeners gather. The prototype of a vowel category works as a perceptual magnet for other members of the category. As a result, participants in Kuhl's study were able to distinguish between instances belonging to

---

[4] Adapted from Chilin Shih: https://netfiles.uiuc.edu/cls/www/course/ling401/week10/categorical.html [Accessed on 18.07.2008]

different categories but failed to recognise the difference between instances within the same category. In that study, participants listened to pairs of vocalic segments, and their task was to decide whether the tokens of each pair were the same or different. The results showed that participants had more difficulty in recognising the difference between sounds that were both close to the prototype, while sounds that were further away from the prototype were less difficult to distinguish. Kuhl suggested that a prototype works as a magnet which draws sounds that are close to it and thus become more difficult to distinguish.

As proposed by Kuhl (2000), infants are born with an ability to differentiate all universal sounds. However, this ability decreases over time until it disappears after a year of age when their production of vowels tightens around good exemplars of vowels in their native language. In this respect, Kuhl, Williams, Lacerda, Stevens and Lindblom (1992) conducted a study comparing English and Swedish infants at the age of six months, with the assumption that their perceptual categories have already been established at this age. These established categories will serve as prototypes and vowels near them will not be differentiated due to the magnet effect. The prototypes used were the English vowel /i/ and the Swedish vowel /y/. Both groups listened to stimuli based on these two vowels. According to the magnet effect, the prototype /y/ would be a magnet for the Swedish infants and the prototype /i/ would be a magnet for the American infants. The results of the study supported the Native Language Magnet model (NLM) proposed by Kuhl. American infants were able to detect differences between variants of a Swedish prototype but were not able to do so with variants of English prototypes. The same was true with Swedish infants who were able to distinguish between non-native variants of prototypes while native ones were not distinguished by them. This study gave evidence that the native language experience had affected the infants' perceptual categories with the result that language-specific categories were developed. Because perceptual categories undergo development, the infants would not be able to make some contrasts that had been possible earlier.

It is obvious from these results that speech perception is affected by language experience from infancy. This perception is shaped by the native language by means of developing native perceptual categories and building up an internal native language system for speakers. This helps speakers of the same native language to discount their individual differences in speech, while on the other hand it may hinder the learning of a non-native language which requires developing new categories for that learned language (Hoopingarner 2004, p.37). Compensating for differences among speakers of the same language related to differences in vocal tract length is a process known as normalisation which is dealt with in section 2.5.1.3 below.

2.5.1.2. Categorical vs. continuous perception

Although categorical perception has been widely recognised as a mechanism of perceiving speech sounds, not all phonological categories seem to be subject to categorical perception. While the perception of consonants tends to be more categorical, perception of vowels has a tendency to be continuous rather than categorical

Liberman, Harris, Hoffman and Griffith (1957) conducted an experiment which involved the use of the plosive+/e/ (as in gate) continuum in which the stop ranged over the /be/-/de/-/ge/ continuum. This was achieved by manipulating the F2 transition in 14 steps to generate the continuum. In previous research (Liberman et al 1954) it was found that the extent and direction of the F2 transition plays a crucial role in the perceived distinction between [d], [b] and [g]. Therefore in generating the stimuli, the starting point of the F2 transition varied in 14 steps of 120 Hz each from 840 Hz below the steady state level in stimuli one to 720 Hz above the steady state level in stimuli 14. F1 transition was left constant in all stimuli because it was found in previous research to be a marker for all voiced plosives (Delattre 1955).

Two types of tasks were conducted: identification and discrimination. In the identification task, all stimuli were randomised and presented to participants who had to label each stimulus with [b], [d] or [g]. In the discrimination task, participants were presented with triads of stimuli (ABX) in which X was either identical with either A or B. Thus, their task was to determine whether X was identical with A or B. The hypothesis was that participants could discriminate stimuli across phoneme boundaries but not within the same category. In both identification and discrimination tasks, participants in the experiment were able to draw clear boundaries between the three consonants.

Another experiment was conducted by Abramson and lisker (1967) and involved the use of /b/-/p/ continua by varying the VOT in incremental steps. These two plosives are produced with the same place of articulation which is the lips; by closing and then opening the lips /ba/ and /pa/ can be articulated. What distinguishes /b/ from /p/ is the time amount between the release of the air and the beginning of the vibration of the vocal folds. This amount of time between the release of air and the vibration of the vocal folds is known as voice onset time (VOT). The VOT is considerably shorter in the case of /b/ than in /p/. In the experiment, participants were asked to identify each stimulus as one of these two sounds. Again, listeners did not have any difficulty in drawing a clear-cut boundary between the two plosive categories when they perceived the stimuli on either side of the boundary as belonging to this or that plosive. These two plosives were not difficult to identify even at the middle of the continua in spite of the gradual change in voicing. The results of these two experiments give evidence for the discontinuous nature of consonant perception or what is known as categorical perception.

While consonants are closer to categorical perception, vowels on the other hand tend to be perceived continuously (e.g., Fry et al 1962, Repp et al 1979). That is, there is no clear-cut boundary between stimuli on a continuum ranging from one vowel to another and in which durational and/or spectral cues are gradually varied. Therefore, there will not be agreement among listeners when perceiving stimuli at the middle of the continuum. Because of the continuous nature of vowel perception, a sort of confusion always exists when the acoustic cues at the middle of the continuum are not sufficient to categorise this stimuli as belonging to this or that vowel. Acoustic cues for vowels are manifest in the formant values they have. For example, a continuum between the vowel /i/ and /ɪ/ can be generated where the former has the formant values of 350 and 1200 for F1 and F2 respectively while the latter has the formant values of 400 and 1800 for F1 and F2. The midpoint of this continuum will have the formant values of 375 and 1600 for F1 and F2 respectively which is supposed to be, theoretically at least, the category boundary between the two vowels. This is so since any stimuli beyond this mid-point will have the formant values that are closer to the vowel on the same side of the continuum rather than the one on the other side. However, listeners will not reach 100% agreement on whether the stimuli close to this midpoint belong to this or that vowel.

Fig.2. 6 An example of continuous perception of vowels taken from one of the experiments in the current study

100%	90%	80%	60%	60%	80%	90%	100%
i	i	i	i	ɪ	ɪ	ɪ	ɪ

As seen in Figure 2.6, the percentage of those perceiving this vowel as belonging to one of the two vowels decreases towards the midpoint of the continuum while it increases towards the ends of this continuum. The formant values at the middle of the continua are not typical of the vowels at the endpoints which represent the best exemplars of these vowels and thus the stimuli in this area are not easy to identify.

This notion of the continuous perception of vowels has been dealt with by a number of researchers. For example, Fry, Abramson, Eimas, and Liberman (1962) investigated the nature of vowel perception by using a vowel continuum of 13 stimuli that ranged from /ɪ/ to /ɛ/ to /æ/. The process of generating the continuum involved manipulating the first two formants. F1 ranged from 330 Hz to 890 Hz and F2 from 1980 to 1760 Hz. Formant frequencies that are typical of the three vowels were obtained from previous work conducted by Petrson and Barney (1952). Eight people participated in the tasks which involved discrimination and identification tests. In the identification task participants were presented with ABX arrangements of stimuli and participants had to label them with the three vowels under investigation. On the other hand the discrimination task involved the use of a forced-choice method whereby participants were presented with triads of ABC in which A and B were always different and X was identical with either A or B. They had to decide which two of these three stimuli were identical and which one was different. Both identification and discrimination tasks showed that participants varied in their recognition of stimuli as belonging to this or that vowel. Therefore, the way in which these vowels were perceived by participants was found to be different from that of plosive consonants in previous experiments. It appeared that there was "no particular sharpening of sensitivity in the regions of phoneme boundaries" (Fry et al 1962, p. 186) in the case of vowels while this sensitivity increased near phoneme boundaries in the case of plosives leading to a rather sharp shift between categories as previous works showed (e.g., Liberman et al 1957). Based on the results of these studies, it was concluded that vowels represent a case of articulatory continuity and thus vowel perception is considered continuous rather than categorical as is the case with consonants.

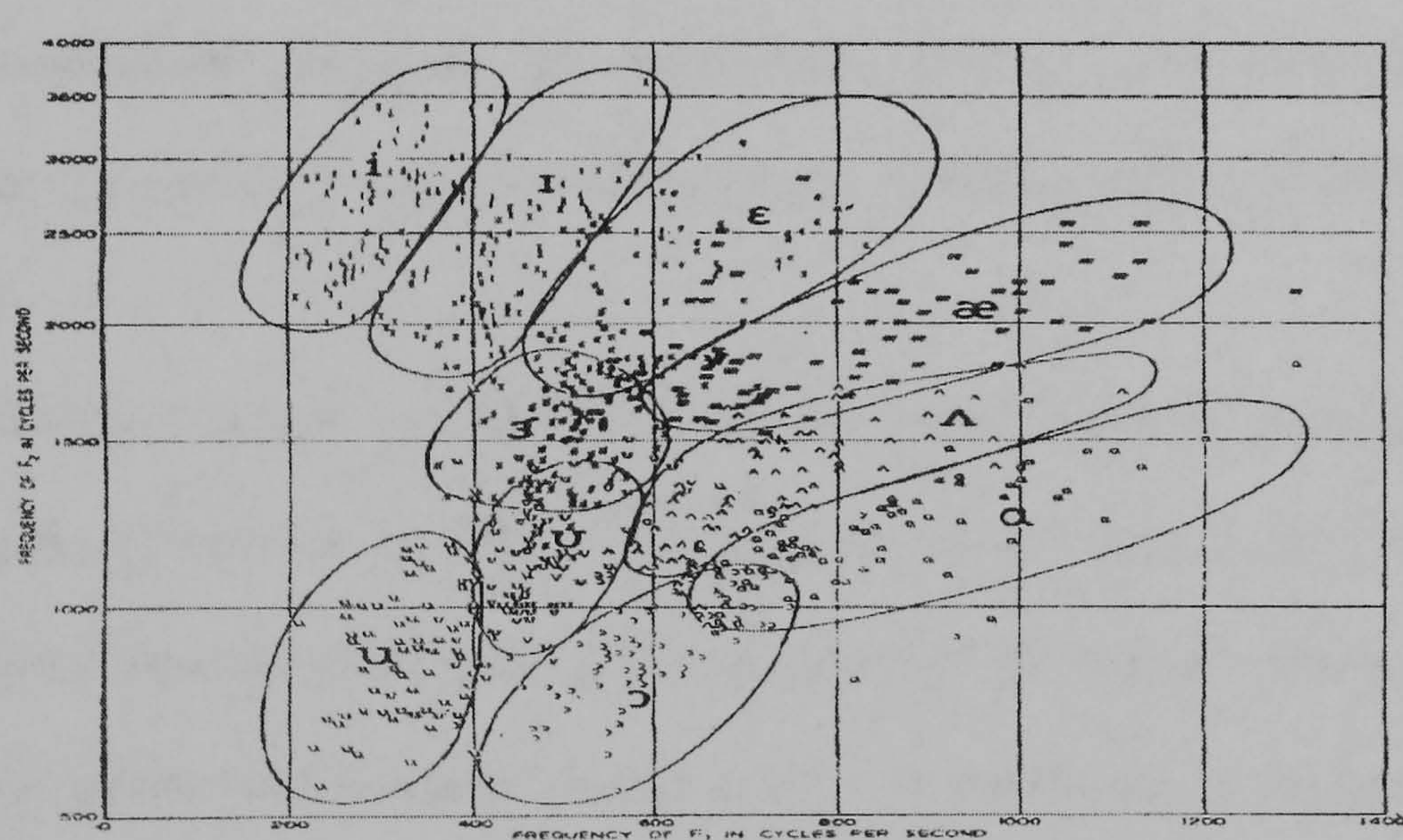
The difference in perception between vowels and consonants in their perception was explained by the fact that consonants have discrete articulatory targets when they are produced while in the production of vowels, the tongue can take different positions when the same vowel is produced. For example, there is a discontinuity between the articulatory configuration required for producing /pa/ and that for producing /ka/. While the former involves movement of the lips, the latter involves raising the tongue towards the palate. However, all vowels involve moving the tongue up and down, and forward and backward in the oral cavity. Therefore, it is possible to change the articulatory configuration of the vowel /i/, for example, to that for the vowel /a/ by gradual and continuous movement of the tongue (with some modifications of the jaw and lip positions).

2.5.1.3. Normalisation

Normalization is a concept that is used to deal with the problem resulting from differences in the vocal tract length. These differences result in differences among speakers in their speech production. For example, female speakers have higher pitch than men, and adults have lower pitch than children because the former have longer and bigger vocal tracts than the latter (see, for example, Petrson and Barney 1952, Fant 1973). When producing vowels, these differences will inevitably lead to differences in formant values between speakers. Therefore, the formant frequencies of a person producing a certain vowel will be different from those of a person producing the same vowel. This may result in some overlap between vowels produced by different speakers having physically different vocal tracts (Ryalls 1996, p.33).

However, the variation between speakers as a result of the differences in vocal tract length is unlikely to cause misunderstanding. The dramatic inter-speaker variation reported by Peterson and Barney (1958) for American vowels (see Figure 2.7 below) did not result in vowel identification problems; all of the vowels represented in Figure 2.7 below were perceived correctly by listeners.

Fig.2. 7 American English vowels variation (Petrson & Barney 1952)



When listeners hear the same utterance produced by different speakers they perceptually eliminate any differences between these utterances and perceive them as one in a process called normalisation. In this process vowels are identified as norms or, as Kuhl (1991) calls them, "prototypes", or "templates" as termed by Joos (1948, p.68). The assumption here is that listeners have an internal system that contains prototypes or templates that they refer to when listening to speech. The sounds they hear are then mapped to these internal prototypes or templates. According to Joos (1948), the vowel space is divided into templates or zones and the vocalic elements that occur within a single zone are perceived as one vowel. This is similar to the 'magnet effect' discussed earlier. Moreover, this internal system is the same in the native speakers of a single language and, thus, normalisation is not likely to cause problems in understanding.

Normalisation is an active process which involves different types of information (Johnson 2005). This information includes visual information about the speakers and other data that help in specifying the speaker's vocal tract size. These include cues external to the perceived vowels such as those derived from the context in which the vowel occurs. An example of this would be the range of formants of the context immediately preceding the vowel. Cues internal to the vowel include, in addition to F1 and F2, F0 and F3 which may help in normalising vocal tract size (Ibid p. 15-16).

However, Speech variation was not only found between speakers but also within them (Johnson et al. 1993). Therefore, the same sound might be produced differently by the same speaker as a result of applying different articulatory strategies. This might indicate that vowel perception is not only dependent on the normalisation of the vocal tract of the speaker. In addition to gender, other differences have also been observed (for example, by Byrd 1994) which may result in phonetic and phonological differences. Cultural factors may also shape differences in speech between men and women. These types of male/female differences may play a role in perception and thus might be incorporated in the normalisation process in order for this process to account for all types of speech variation.

Nusbaum & Morin (1992) classify theories of speaker normalisation into two types: contextual tuning theories and structural estimation theories. Contextual tuning theories maintain that a listener uses a sample of speech to learn the vocal characteristics of a speaker and uses these characteristics to interpret the phonetic structure of subsequent utterances. On the other hand, structural estimation theories believe that each sample of speech is normalised independently. Listeners use information that is internal to the structure of an utterance to estimate the vocal characteristics of the speaker (Nusbaum & Morin 1992, p. 114).

However, according to Nusbaum and Morin, listeners use both of these mechanisms of structural estimation and contextual tuning depending on the degree of speaker variation. Structural estimation is used when there is variation in speakers from utterance to utterance. In this kind of normalisation, the listener benefits from information within the utterance to understand the context necessary for identifying that utterance. In vowel perception, for example, listeners make use of F0 and F3 besides F1 and F2 in order to recognise the vowel. However, when the speaker is constant, the listener uses the contextual tuning mechanism in which acoustic-phonetic relationships among sets of segments and utterances are exploited rather than depending on a certain number of cues contained in one utterance (Nusbaum & Morin 1992, p.131).

Despite efforts to address the problems of invariance and linearity within auditory-based theories, the search for such invariant cues for phonemes has remained problematic due to the variability that characterises the acoustic signal. It has become clear for researchers that phonemes are not discrete elements when they are in sequence in the speech waveform (Fant, 1973). Rather, they are affected by several factors when they occur in context, which, in turn, affect their physical realisation and their acoustic cues. As Cooper et al (1952) put it:

> ...the important point, however phrased, is a caution that one may not always be able to find the phoneme in speech wave, because it may not exist there in free form; in other words, one should not expect always to be able to find acoustic invariants for the individual phoneme (Cooper et al 1952, p.605).

An alternative model to the vocal tract normalisation is the exemplar approach to speaker normalisation. In this model, vowel categories are represented by all instances of the vowel experienced by the listener rather than merely by abstract representation as in the vocal tract size normalisation. According to this approach to speaker normalisation, the reference system is the set of all examples of the vowel experienced by the listener rather than a single representative abstract prototype of that vowel. In such normalisation exemplars that are suitable are activated and those that are not are deactivated (Nosofsky 1988).

The advantage of the exemplar approach is that it accounts for all types of variation mentioned above and not only those resulting from vocal tract size. Linguistic categories are no longer seen as invariant abstract representations. Rather, they are sets of all instances of the sound experienced by the speaker/listener of the language (Johnson 2005).

2.5.2. Motor theory

Motor Theory (henceforth MT) suggests that listeners interpret speech sounds in terms of the motoric gestures they would use to produce those sounds (Liberman et al 1967). Therefore, instead of looking for invariance in the acoustic signal, the listener directly perceives speech gestures as intended by the speaker (Fodor 1983).

This theory is based on the assumption that perception is strongly related to production by a number of neural representations. Those who proposed this theory argue that "[t]hough we cannot exclude the possibility that a purely auditory decoder exists, we find it more plausible to assume that speech is perceived by processes that are also involved in its production" (Liberman et al., 1967, p. 452).

To solve the problem of variability resulting from coarticulation, MT suggests that the listener, having knowledge of how sounds are produced in isolation and how they are coarticulated, can manipulate this knowledge to decode the speech signal into appropriate phonological elements (Liberman & Mattingly 1985, p.6). Some other researchers go even further, claiming that phonological elements coincide with articulatory gestures. As Fowler and Galantucci (2005) put it:

> [P]honological forms are in fact the public actions in which speakers engage when they talk, and consequently, they may be preserved throughout a communicative exchange. They are the atoms of talkers' plans to speak and of their vocal tract activity (Fowler & Galantucci, 2005, p.637).

However, further investigation (e.g., Mattingly & Liberman 1969, Liberman 1982) suggested that the perception process proposed by the theory was not sufficient. It was found that the articulations of the speaker from which the listener retrieves the message are not invariant. In a revised version of motor theory (Liberman & Mattingly 1985), and in order for the theory to overcome the problem of variability in articulatory gestures, it was proposed that the linguistic representations of speech are abstract units in the mind of the speaker and, thus, listeners perceive the intended articulatory gestures while relying on their knowledge as speakers. The phonological segments that are used to convey the message to the listener exist as components of the speaker's knowledge, and for the listener to perceive that phonological message these components should be part of their knowledge as well (Fowler & Galantucci, 2005, p.635). This knowledge enables the listener to recreate the intended phonemes through an internal speech synthesis process by which coarticulatory effects are eliminated (Rosner & Pickering 1994, p.372).

One of the advantages of MT is that it presumes a link between the production and perception of speech by claiming that "speech perception is not to be explained by principles that apply to perception of sounds in general, but must rather be seen as a specialisation for phonetic gestures" (Liberman & Mattingly 1985, p.6).

However, although MT was meant to address the lack of invariance problem, it has not provided suggestions for dealing with speaker variability (Nusbaum 1992, p.113). Moreover, the question of how listeners gain access to gestures intended by speakers remains unanswered. It is not yet clear how the listener makes a link between what they hear through the acoustic signal and the motor commands or neural representations of speech in the speaker's mind (Ryalls, 1996, p. 59). Proponents of the theory argue that there must be some innate structures in the brain that directly decode acoustic information into intended gestures. This innately specified link between production and perception is evident in pre-linguistic infants who are able to categorise possibly all phonetic distinctions. It is only after a year of age when these infants begin to lose sensitivity to linguistic gestures that they do not use due to the process of acquiring their mother tongue (Ryalls 1996, p.24). However, being unique to the human being, these structures cannot be investigated.

2.5.3. Direct Realist Theory

Direct Realist Theory (henceforth DRT) was developed by Fowler (1986), and is based on the assumption that the perceiver relies not on the sensory stimuli but rather on environmental events. Those events that are relevant to speech perception are called "distal events" (i.e. the articulations). It is the distal object or the phonetic gesture of speech production that is perceived, and the acoustic signal is only a medium for providing information to the perceiver about that event. As Stetson (1951; cited in Lindblom, 1996, p.1691) puts it, "Speech is rather a set of movements made audible than a set of sounds produced by movements". The proponents of this theory often draw analogy to visual perception. When we see, we do not see the light reflected from objects but we see the objects themselves. The same happens with speech perception. That is, what we hear is the articulatory sounds and not the sound waves that carry them.

According to Fowler (1986), listeners seek information about events in the environment from which they disentangle separate phonemic articulatory gestures. Variation in the speech signal is seen by the theory as a rich source of meaningful information about the event, including the dialect, gender and emotional state of the speaker, which are crucial for perception. Variation is no longer seen as noise that causes a problem to the listener and that should be eliminated through a process of normalisation (Wright et al., 1997, p. 29).

However, as is the case with MT, DRT does not make it clear how listeners use the acoustic stimulus to perceive articulatory gestures. As Klatt (1988) describes them, these theories are plausible descriptions but they are difficult to test since they are computationally empty, or as Ohala (1996, p.1719) puts it, these theories have so far failed to provide an algorithm for deriving the articulatory gestures which generate the speech signal.

Because of the conflicting views held by these theories, some researchers have tried to look for a compromise between articulatory and acoustic auditory perspectives on perception. In this respect, Nearey (1997) suggested what he calls the double-weak approach, as opposed to the strong theories mentioned above. This approach, like MT, maintains the view that phonological forms originate in natural articulatory processes. However, it also sides with the AIT by holding the view that "the real-time objects of perception are well-defined auditory patterns" (Nearey 1997, p. 3241). Nearey's argument here is that since strong theories of perception entail assumptions that are too strong to be evidenced in the light of the data available, it is better to adopt an approach that combines weaker versions of the assumptions of both types of theories. According to this double weak approach, speech perception requires two conditions to be met. First, symbols are to be encoded into articulatory gestures. Then, the acoustic signal representing these articulatory actions should provide sufficient auditory cues that enable the listener to perceive the intended symbols through a process of decoding (Neary 1997, p.3243). This approach is in some respects similar to the Fuzzy Logical Model of perception reviewed in the following section.

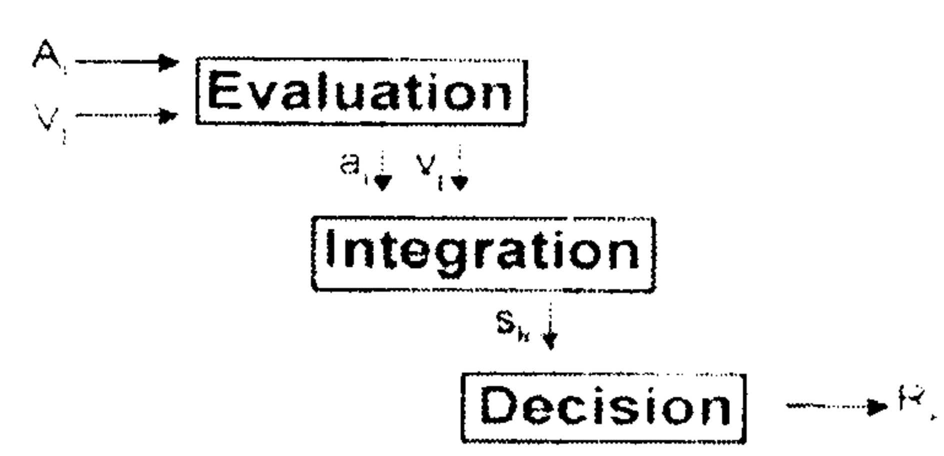
2.5.4. Fuzzy Logical Model of Perception

The Fuzzy Logical Model of Perception (henceforth FLMP) (Oden & Massaro 1978, Massaro & Cohen 1983, Massaro 1989) aims to provide a mechanism for integrating the information necessary for perception from different sources, including auditory and visual input. This is important considering the role that can be played by visual input in speech perception. People often obtain information from multiple sources, since a single source might not be enough to specify the appropriate interpretation.

The idea of prototypes is crucial in this model of perception. A prototype is a category that has certain features, which has a representation in memory and a sensory (auditory and visual) representation. These two types of representation must be compatible and related to each other in order for appropriate perception to be fulfilled. For example, to identify the syllable /ba/, the listener must make a connection between this syllable and the memory of the category /ba/ (Massaro 1992, p.80). The FLMP considers syllables as the fundamental perceptual units of speech perception.

This model assumes that there are several sources of information which support speech perception. According to this model, information processing for perception involves three stages: evaluation, integration, and decision (Massaro 1992, p.79) as shown in Figure 2.8 below.

Fig.2. 8 Speech perception stages (Massaro 1992, p. 79)



Although these processes are sequential in time, they also overlap. In Figure 2.8, sources of information are represented by upper-case letters. A_i stands for auditory information and V_j refers to visual information. The evaluation process changes these sources of information into psychological values (referred to by lower case letters a_i and v_j which are evaluated independently to assess the alternatives provided by each information source. In other words, this stage involves analysing the features of the input signal. In the second stage, the different sources of information are integrated to provide an overall degree of support S for each speech alternative k. This integration involves comparing all phonological features that contribute to the final value of the perceived element with stored prototypes, and the feature most compatible with the stored prototype will have the most influence on the outcome. In the final stage the output of integration will be mapped into some response alternative R_k and a decision made to choose the best fitting prototype to fulfil perceptual identification and the interpretation of the language input (Massaro 2001, p.3). This decision is affected by factors such as how many times the chosen syllable was identified as an example of a prototype, the degree of matching between the exemplar and the prototype, and the degree of information provided by a feature in that the most informative feature will have the most influence in making a choice.

Evidence for the integration of sources of information in the perception process comes from the so-called 'McGurk effect' (McGurk & MacDoland 1976, p.746). Here it was found that listeners combine visual (articulatory) information with auditory information in order to perceive the sound. In an experiment, listeners were shown a person's face visually producing [ga] while simultaneously hearing [ba]. Listeners reported hearing [da]. In effect, listeners exploited both what they saw and what they heard, and perceived an articulation somewhere between the two.

However, because the FLMP approach is based on mapping features to prototypes stored in the memory, it will not be possible to account for all speech variations. Moreover, the connection between speech production and perception is also missing (Wright et al., 1997, p. 31).

2.5.5. Recent developments

One aspect in common in all the theories discussed so far is their emphasis on the early abstraction of the signal to phonological units such as features, phonemes, or gestures during the perception process, regardless of whether these units are auditory or articulatory in nature. This is mainly due to the search for simplicity in phonological representation and the neglect of phonetic detail as irrelevant noise. Since the early 1990s, researchers have come to the conclusion that, whatever the units of perception are, they cannot be separated from the context in which they occur and that fine phonetic detail can be very informative from a linguistic point of view. However, in recent research into speech perception, variability in the acoustic signal is no longer seen as a problem. It is instead considered as something that enhances perception due to the rich linguistic and non-linguistic information it contains (Goldinger 1998). The exemplar approach, for instance, considers all instances of a vowel experienced by a listener representations of the category of that vowel. According to this approach all variations of a vowel category are accounted for and not merely those resulting from differences in vocal tract size (see section 2.5.1.3).

Linguistic cues for vowels, for example, may include their quality represented by their formant values and their quantity. On the other hand, non-linguistic information includes the emotional state, the social context and the speaking rate, among other things (Pisoni 1997) that the listener exploits during speech perception. Recent work has shown that phonetic detail can facilitate the identification of words in connected speech (e.g. Smith 2004) and can aid native language acquisition (e.g. Best 1995). How this phonetic detail is stored in the brain and at what stage abstraction occurs remain, however, the subjects of much debate.

2.5.6. Concluding remarks

Theories of speech perception have addressed the issue of what is being perceived by looking at the articulatory, visual and auditory domains. Whether the perceived components are articulatory gestures, phonetic features, phonemes or syllables also varies among these theories. Although they differ in their means, most of the theories mentioned "assume that, once the right way of treating the phonetic measurements has been found, variability will be significantly reduced and will thus cease to be a problem" (Lindblom, 1996, p. 1684). That is, most theories believe that speaking is controlled by underlying invariant goals. However, what these goals are and how they can be recovered by listeners is still a controversial issue. While these goals are articulatory in some theories, as for example with MT, they are acoustic/auditory with others such as AIT.

However, the question that arises is whether invariance is necessary at the articulatory or the acoustic/auditory level. According to Lindblom, speech perception does not presuppose the existence of invariance at these levels. Invariance should only be present in the final goal of the speaker which is the lexical item intended by the speaker. All pronunciations of the same word, however much they differ, lead to the recognition of that word. "[T]he essence of the speech production process is ... an elegantly controlled variability of response to the demand for a relatively constant end" (MacNeilage, 1970, p. 184).

Moreover, the speech signal does not necessarily have to contain all the information needed for perception (Lindblom, 1996, p. 1687). In fact, speech perception results from the interaction of information contained in the speech signal at all its articulatory, acoustic and auditory levels with the phonetic, lexical, grammatical and other stored knowledge of the listener. The role of all of these levels of representation is to facilitate the perception process. The process of lexical discrimination is facilitated by what Lindblom calls signal-complementary processes or knowledge, a thing which is external to the signal itself, and that the role of these complementary processes and their contribution in signalling a contrast is estimated by the speaker (Lindblom 1990, p. 404-5).

Due to the complexity of speech perception, theories to explain it are not expected to have strong predictive power. As Massaro points out, theories are expected to have predictive power only in laboratories where the complexities of natural settings are simplified, controlled and measured (Massaro 1992, p. 82). However, the different theories discussed above and the controversial issues resulting from them should not discourage researchers from further investigation in the area of any of these three levels of speech signal representation. On the contrary, such articulatory, acoustic or auditory studies, will provide better understanding of how humans communicate, give more insights into the field of speech, and help in solving controversial issues such as those related to vowel perception.

2.6. Approaches to vowel perception

Perceiving a vowel involves two processes: identification and categorisation. The former process is related to identifying the same vowel when produced in different contexts and by different speakers. The latter, however, is concerned with distinguishing vowels from each other (Rosner & Pickering, 1994, p. 1). In terms of vowel categorisation, it is widely believed that the first two formants are the primary acoustic cues which are used by listeners when perceiving vowels (Raphael, 2005, p. 197). These formants may show considerable variability depending on factors such as differences in vocal tract length between speakers, speaking rate, consonantal context and stress, all of which may influence vowel identification. In spite of this variability, listeners are still able to recognise different sets of formants as the same vowel through a normalisation process that enables those listeners to compensate for such variability.

More research on vowel acoustics has shown that higher formants, i.e. F3 and above, and the fundamental frequency (f0) may also play a role in vowel perception. In a study conducted by Fujisaki and Kawashima (1968) it was possible to draw boundaries between synthesized vowels by merely varying f0 and F3. For example, it was found that when f0 is shifted by 200 Hz, F1 also shifts by 100 Hz to 200 Hz. On the other hand, an F3 shift of 1500 Hz results in a category boundary shift of 200Hz in the F1-F2 plane for the /u-e/ continuum and 50 Hz for the /o-a/ continuum. Other work has also highlighted the importance of formants other than F1 and F2 in vowel perception (see, for example, Nearey 1989, Nusbaum & Morin 1992, Halberstam 1998).

Another factor involved in the perception of vowels is their duration. Although duration is considered a secondary cue in languages such as English, where listeners mainly depend on the quality of the vowel as the main cue for perception, the role that duration may play in the production of languages like Arabic leads one to predict that it may play an equally important role in perception. This study investigates this role in the context of LA.

Finally, the context in which the vowel occurs can also affect the perception of that vowel. In an experiment conducted by Ladefoged and Broadbent (1957), it was found that the identification of a vowel is affected by the formant frequencies used in the preceding sentence. Therefore, the perception of a vowel in a /bVt/ syllable following the sentence 'Please say what this word is' varied when the formant frequencies used in this sentence varied. It was found that when vowel formants contained in the precursor phrase are shifted up, the following vowel in a /bVt/

syllable is perceived differently from when these formants are shifted down. This means that the precursor phrase, though external to the vowel, provides a coordinate system for the identification of that vowel. This may indicate that vowel identification is not only dependent on characteristics internal to the vowel but also involves external factors.

The main question related to constructing a theory of vowel perception has been concerned with the nature of vowels as articulatory or acoustic events and how these vowels are perceived by listeners as they are intended by speakers. In this respect, there are two major approaches which deal with these questions about the characterisation of vowels. These are the "static target" approach and the "dynamic specification" approach (Strange 1989, p.2082).

2.6.1. Static target approach

The static target approach is based on the notion of "vowel targets" (Strange 1989, p.2081). Vowels have been traditionally contrasted on the basis of the static vocal tract positions taken by the tongue, the jaw and the lips. The acoustic patterns resulting from these articulatory targets are referred to as acoustic targets. Therefore, vowel targets are seen as the unaffected stored representations of vocalic segments (Daniloff & Hammeberg 1973, p. 240). Because these targets are seen as static points of articulation, they can be sufficiently represented as points in the vowel space coordinated by the first two formants. These two formants are often adequate for the perception of these vowels (Delattre et al 1952, p.196). However, considering vowels as static targets often encounters two problems. As already been mentioned, target vowel formants are often ambiguous because of inter-speaker variation which leads to overlapping in the vowel space. Another source of problems is coarticulation. When vowels are produced in the context of consonants, they are often "undershot" and, therefore, the target is often not reached (Lindblom 1963, Gay 1978).

In order for the listener to perceive vowels, they need to compensate for such variability resulting from these two problems. As has already been noted above, this is achieved through a process known as normalisation. The aim of the static target approach is to account for speaker normalisation Strange 1989, p. 2083). According to this approach, the steady-state portion of the vowel is considered the most essential part which represents the vowel target and by which the vowel can be identified. However, as said before, vowels vary according to speaker and context. In normal speech, the steady state of the vowel might not be reached. Yet, listeners are still capable of identifying vowels. Therefore, how normalisation (see section 2.7.3) is achieved is still a major issue in the effort to construct an adequate theory of speech perception.

2.6.2. Dynamic specification approach

As a reaction to the target vowel approach which has treated the vowels as static points in the vowel space, the "dynamic specification" approach emerged due to the realisation among some researchers (e.g. Bladon 1985, Strange 1989) of the great importance of dynamic aspects of speech. This approach was motivated by findings reported by Peterson & Lehiste (1960) which suggested that vowels can be perceived not as static targets or points in the vowel space but as dynamic acoustic patterns.

A study on American English vowels revealed that there were systematic differences between vowels in both duration and in the shape of formant trajectories or transitions from and to the syllable nucleus. For example, it was found that the duration of all English vowels is significantly affected by the nature of the following consonants in that vowels were shorter before voiceless consonants and longer before voiced ones.

Moreover, consonant class also has an effect on vowel duration. Plosives were found to be preceded by the shortest vowels while those followed by fricatives were the longest. In general, it was found that American English vowels can be classified in two groups in terms of their duration: short vowels [ɪ, ɛ, ə, ʊ] and long ones [i, eɪ, æ, ɔ, ou, u, ɑu, ɑɪ, ɔɪ] (Peterson et al 1960, p.11). Systematic differences in formant trajectories were also found between tense and lax vowels. Syllables containing tense vowels had more rapid transitions into and slower transitions out of the syllable nucleus as well as much longer quasi-steady-state portions than those with lax vowels. These differences in dynamic acoustic patterns have been considered to be perceptually important because they distinguish vowels that overlap in the vowel space due to articulatory and acoustic similarity (Strange 1989, p. 2082).

Another study by Lindblom and Studdert-Kennedy (1967) also found that the identification of a vowel does not only involve the formant pattern at the midpoint but also the direction and rate of adjacent formant transitions. For example, listeners were found to draw a boundary in the /ɪ/-/ʊ/ continuum differently when the vowels were in a /w-w/ context from when they were in a / j-j/ context.

In harmony with this dynamic approach, Strange (1989) argues that there is dynamic information available in coarticulated syllables that listeners utilise to identify the speaker's intended vowels. Besides the central steady-state portion of the vowel which provides essential information for its recognition, there are other sources of information. These include the relative durational differences intrinsic to the vowel and formant transitions into and out of the syllable nucleus (Strange 1989, p. 2084; Strange et al, 1983, p. 703).

Other acoustic studies (Harrington & Cassidy 1994, Huang 1992, Zahorian & Jaghrghi 1993, Hillenbrand et al 1995) have also demonstrated that vowels are more effectively distinguished if information related to acoustic parameters is extracted from different temporal points rather than only from the mid-point in the vowel target. For example, formant transitions into and out of the quasi-steady-state of long vowels in context tended to be equal in duration and slope while those for a short vowel were found to be different. That is, formant transitions out of short vowels are found to be more gradual and longer than those into it (Strange 1983, p. 704).

Watson and Harrington (1999) conducted a study on Australian English vowels to assess the effect of formant trajectories on vowel recognition. The results showed that vowels differ in the extent of dynamic information relevant to their identification. Monophthongs can be equally classified using static or dynamic information whereas diphthongs are more accurately identified relying on dynamic information because this type of vowel is characterised by two targets rather than one as in the case of monophthongs. However, dynamic information was found to be useful in distinguishing between tense monophthongs and their lax counterparts, because the former have proportionally delayed targets compared to those of the latter (Watson & Harrington 1999, p.466).

As Strange (1989) suggests, this shift from the focus on vowels as static points to their dynamic characterisation has implications for solving the problem of invariance that has long hindered efforts to construct an adequate theory of speech perception. If the information included in a single syllable is enough to identify the vowel included even when such vowels are affected by coarticulation with neighbouring consonants, the aim will be to develop a perceptual model that describes how listeners detect these different sources of information and utilise them to identify the vowel intended by the speaker (Strange 1989, p.2085).

2.7. Summary

Speech production and speech perception are intimately linked to each other. Together they constitute the speech chain which starts from the brain of the speaker and ends in the brain of the listener. Speech is marked with variability due to several factors including inter and intra-speaker variation. As far as vowel inventory is concerned, variability is also found across languages, which differ in inventory size and inventory vowel space.

There exist a number of theories that try to explain this variability in production and how this speech variability problem is dealt with in perception. Theories of speech production include QT and SFT, and those related to perception include AIT, MT, DRT and FLMP.

Speech variability also affects the way vowels are perceived. Listeners often resort to normalisation and categorical perception in dealing with such speech variability. In this respect, two major approaches to vowel perception exist, namely the static target approach and the dynamic specification approach.

The aim of this chapter was to offer a theoretical framework for this study that may help in developing research ideas and finding explanations for the results obtained. This would also help in determining the appropriate methods to deal with the production and perception data in this study and the factors that should be considered. The next chapter reviews the literature related to the Arabic vocalic system. The review will also cover studies that have dealt with the Libyan Arabic dialect and particularly with its vocalic system.

CHAPTER III

THE ARABIC VOCALIC SYSTEM

3.0. Introduction

Arabic is a Semitic language spoken in what is commonly known as the Arab World. The region stretches from the Arabian Gulf in the east to the Atlantic Ocean in the west. This includes all countries in the Arabian Peninsula and North Africa. It is considered the sixth most widely spoken language in the world (Newman 2002, p.1). The language is also considered important for Muslims. The holy Koran was revealed in Arabic, and Muslims believe that it should be read in Arabic in order for them to understand the message of God (Huthaily 2003, p.1). Therefore, many people in Muslim countries are interested in learning the language and speaking it. The language has also gained interest in the whole world because of the increasing political and economic importance of the Arab world.

The variety of Arabic known as Classical Arabic (henceforth CA) was originally spoken in Mecca (in Saudi Arabia) where the Prophet Mohammed was born and it is the variety by which the holy Koran was revealed in the seventh century AD. At present, the use of this variety has become very limited and the only domain where it is still used is in recitations of the holy Koran. The dominant variety nowadays in the Arab world is known as Modern Standard Arabic (henceforth MSA), the official language in the Arab countries and the language used in educational institutes and the media. The major differences between CA and MSA are that the latter has a larger vocabulary and uses less complicated forms of grammar (Huthaily 2003, pp.1-2).

Apart from the standard variety of Arabic, there exist dialectal varieties that are used in everyday speech and informal situations. The coexistence of a normative variety and a vernacular variety, where the former is mainly written and used formally whereas the latter is mainly spoken and used informally, is a situation known as diglossia, in which both varieties play complementary roles (Wardhaugh 2002, p.88). The dialectal varieties of Arabic do not only differ from one Arabic speaking country to another but also from one region to another in the same country. The term dialect refers not only to differences in pronunciation but also to grammatical and vocabulary features which distinguish one dialect spoken in a certain region from another spoken in a different region. In Libya, for example, the dialect spoken in Tripoli (i.e. Tripolitanian Arabic, henceforth TA), the capital city which is in the north west of Libya, is different from that spoken in Derna, (i.e. Derna Arabic, henceforth DA) a city in the north east of the country. These different dialects also differ in accent. The following words, for example, are pronounced differently in the two cities (Aurayieth 1982, p.3).

(3.1)	MSA	TA	DA	Gloss.
	kataba	ktab	kitab	he wrote
	fahimat	fihmit	ʔifhimat	she understood
	banaat	bnaat	binaat	girls

In addition to the dialectal differences which range from vocabulary to grammar, these accentual differences may result in problems of intelligibility between speakers from different dialects (Al-Ani 1970, p.18; Newman 2002, p.1; Alghamdi 1998, p.4). Moreover, the phonetic implementation of the MSA vowel system differs according to dialect (Alghamdi 1998, p.8). This means that the MSA spoken by Libyans is marked by Libyan dialectal features. In this study the focus is on the vowel system of the Libyan dialect, the variety spoken in Libya, since there are very few descriptions of LA vowels. However, this dialect shares some features with other Arabic dialects considering their common origin and taking into account the fact that all speakers of these dialects come from countries where the same standard variety is used officially. It is therefore necessary that these various dialects and the features they share, especially with regard to vowels which are the main emphasis of this study, should be considered in this literature review.

3.1. Arabic vowel inventory

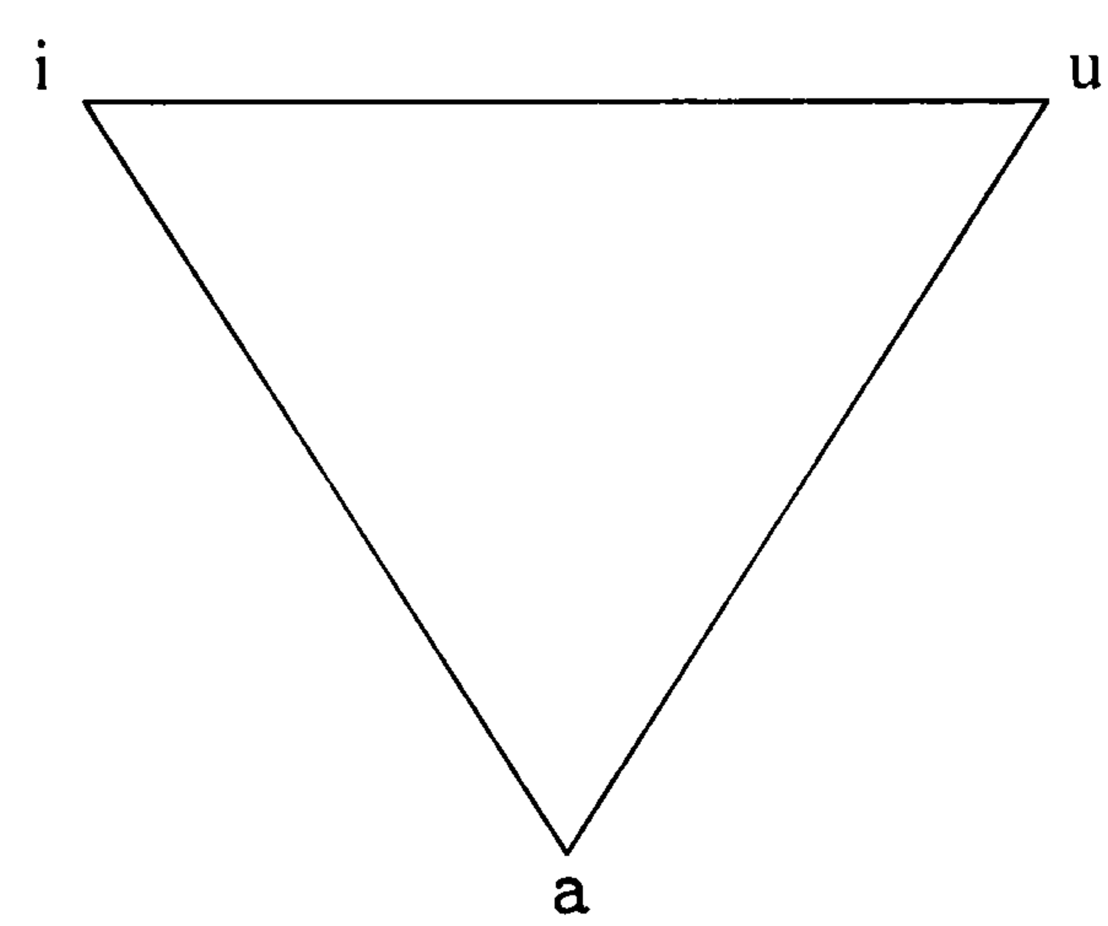
Compared with languages such as English, MSA has a small number of vocalic sounds. Only three main vowel qualities can be distinguished: low central, high front and high back. These three main qualities are the most common in the languages of the world and are known as basic or fundamental qualities (Maddieson 1984). However, these vowel qualities are realized in two forms each: short and long (Mitchell 1993, p.138). In addition to these vowels which are found in both the standard variety of Arabic and almost all Arabic dialects (Alghamdi 1998, p.4), there are some vowels which are unique to certain regional dialects. The dialectal vocalic sounds which are found in the Libyan dialect are discussed in a separate section.

Newman (2002) conducted a study that aimed at finding out how unique the Arabic phonemic inventory is within the world's languages. The results were mainly based on information taken from a database that belongs to the University of California, Los Angeles (UCLA) known as the Phonological Segment Inventory Database (UPSID). The research showed that the Arabic vowel inventory is well below the mean within the UPSID languages in terms of having only three main vowel qualities, /i/, /u/ and /a/. However, these three vowel qualities were found to be the most common, being respectively found in 91%, 83.9% and 88% of the languages included in the database (Newman 2002, p.9).

3.1.1. Vowel quality

The MSA vocalic system contains three fundamental vowels. These three fundamental vowels in Arabic are often represented in the literature by "a sort of inverted triangle, with apex low at ɑ, and raised base i-u" (Gairdner 1925, p.33) as shown in Figure 3.1. The fact that the Arabic vowel system is a simple one which contains only three main vocalic units was also reported by Mitchell (1993, p.138).

Fig.3. 1 Fundamental vowels in Arabic

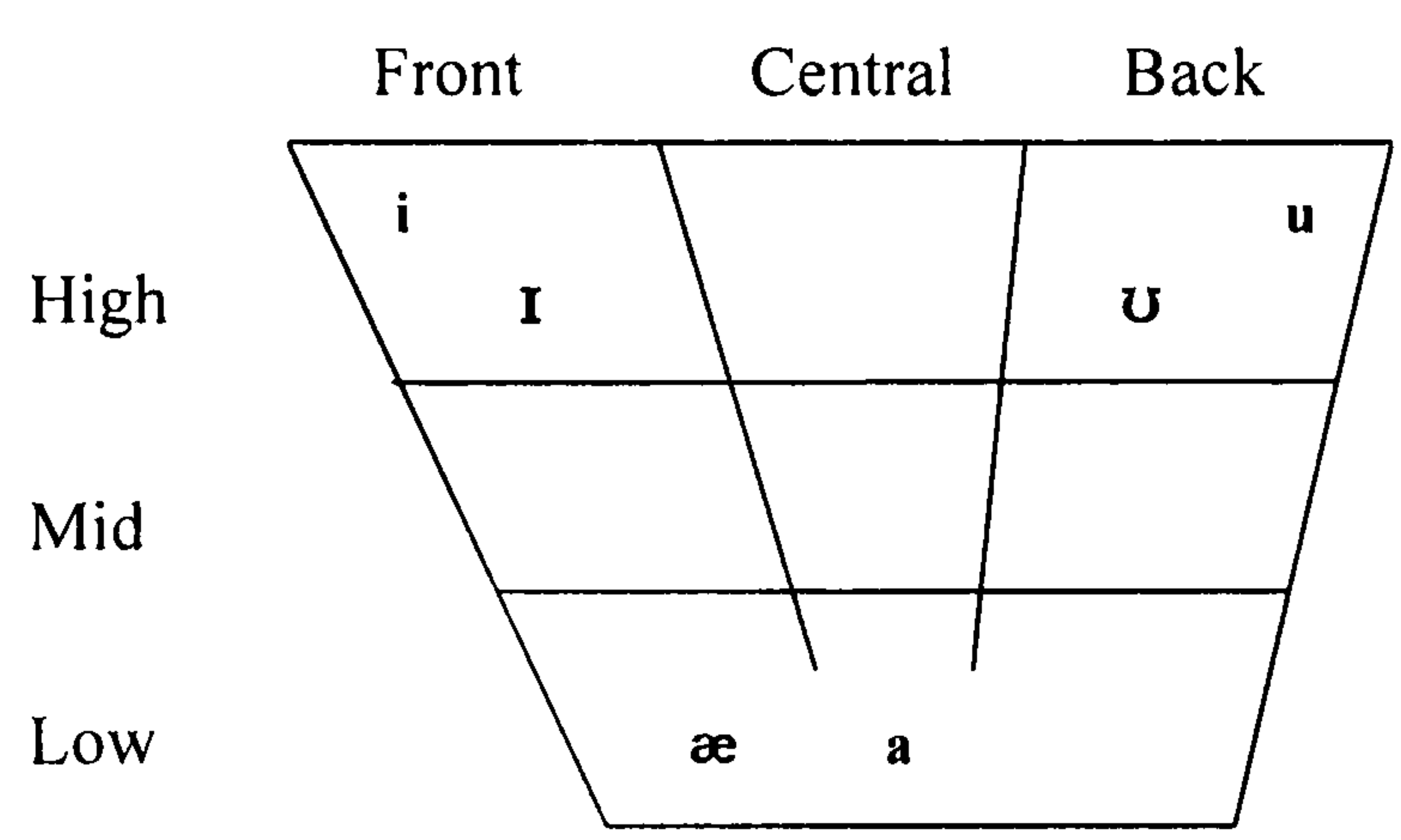


As can be noticed here, the symbol used by Gairdner for the low vowel which is /ɑ/ is replaced in the triangle with /a/. As will become clearer later, the symbol /ɑ/ is only used to represent the emphatic version of the vowel, which is further back and lower than the plain version of the same vowel. It is more convenient to use /a/ as a phonemic symbol that includes both plain and emphatic allophonic variations.

This traditional view that Arabic has only three main vocalic qualities which are realized in a short and a long form each is now disputed by some researchers.

Huthaily (2003, pp.28-30), for example, refers to the long vowels as /i, u, a/ but symbolizes their short counterparts using /ɪ, ʊ, æ/ respectively. This is reflected in the way he distributes the short and long forms of each vowel in different positions in the vowel quadrangle, which indicates the quality difference between the two elements of each pair. For example, the short form of the low central vowel /a/ is located in a more frontal position and is symbolised as /æ/ as shown in Figure 3.2 below.

Fig.3. 2 Approximate locations of MSA simple vowels (Huthaily, 2003, p.30)



However, Huthaily's description of these vowels was mainly theoretical and is based on impressionistic speculation since no reference was made to any experimental work upon which these descriptions can be based.

Gairdner (1925, p.37) had a similar view about the short version of /a/ with regard to its location, though he refers to it using a different phonetic symbol which is /ɑ/. He maintained that this vowel resembles the IPA [ʌ] notation in its articulation and that its long version /ɑː/ is, in fact, produced as in the IPA [æ]. His description was not supported by robust data. My intuition as a native speaker of the language is that the IPA /ʌ/ is only similar to the emphatic short version of /ɑ/, which is only found in the vicinity of emphatic consonants (see section 3.1.1.1 below for a full review of the effect of emphatic consonants on vowels).

Cowan (1970, p.96) is another proponent of the use of /ɪ, ʊ, æ/ to represent short vowels. He further claims that the use of the length mark (ː) after long vowels is redundant and considers it as merely an allophonic feature of these vowels. According to Cowan (1970), length is not the only factor that distinguishes these vowels. However, while duration is not the only distinguishing factor between tense and lax vowels in languages like English, phonological length plays a major role in others, such as Arabic.

The uses of different symbols for short vowels from those used for long ones by Gairdner (1925), Cowan (1970) and Huthaily (2003) are all based on impressionistic speculation. Therefore, experimental and objective proof for the qualitative differentiation between short vowels and their long counterparts and, subsequently, for the use of different symbols for them, is still lacking. It is the main aim of this study to investigate these vowels acoustically and auditorily in the Libyan variety of Arabic in order to find out how the quality and quantity of LA vowels interact and to what extent quality participates in the perceptual distinction between a short vowel and its long counterpart.

This is not to say, however, that relevant acoustic studies have not been conducted. Some studies of Arabic dialects have included acoustic analysis. The first acoustic study of Arabic vowels was conducted by Al-Ani (1970). His work aimed at studying the phonology of the standard variety of Arabic used in Iraq. The description of vowels by Al-Ani was based on articulatory and acoustic data gathered from Iraqi informants which was analyzed using X-ray tracings and spectrographic displays. However, although vowels were recorded in isolation as well as in minimal pairs and sentences, the description of vowel formants and duration relied mainly on the recordings made of isolated vowels produced by the researcher himself. Al-Ani justified this by arguing that consonantal context may result in vowel consonant transitions which may affect the steady state formants of vowels. This resulted in vowel durations being twice as long when isolated as when they were spoken in utterances (Al-Ani 1970, p.22) (see section 2.1.1.3 on vowel length), which may have affected the realisation of these vowels both in terms of the researcher being aware of the purpose of the study and the unusual length of the vowels. Table 3.1 shows the average formant frequencies of Arabic vowels as obtained by Al-Ani (1970, p.23).

Table 3. 1 Means of formant frequencies of Arabic vowels (Al-Ani 1970, p.23)

i:		i		u:		u		a:		a	
F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
285	2200	290	2200	285	775	290	800	675	1200	600	1500

It is obvious from Table 3 above that there is little difference between the long high vowels /i:/ and /u:/ and their short counterparts /i/ and /u/. However, this difference is greater between the long low vowel /a:/ and its short counterpart /a/; Al-Ani suggested that the former is more retracted and lower as evident in its higher F1 and lower F2.

More recently, a formant frequency analysis study of Arabic vowels in connected speech was conducted by Newman & Verhoeven (2002). The Arabic style that was chosen for investigation was the Koranic recitation style because of the prestige and purity of this variety. However, vowels in pharyngealized context were excluded due to coarticulation effects. The material analyzed included 30 minutes of recorded recitation by one of the most respected Koran readers among Arab-Muslims. This recorded sample included 400 vocalic observations covering the vowel categories found in this language variety. Along with this investigation into Koranic recitation style, an acoustic analysis by the two researchers of the same vowels in connected colloquial Egyptian Arabic was also conducted. A translated passage from English to Arabic was read by a speaker of Egyptian Arabic and recorded. The choice of this dialect was attributed to its prestige among other dialects in the Arab world. The data gathered were analysed acoustically. Table 3.2 shows the average formant values of the vowels investigated in both Classical Arabic and Cairene Arabic (Newman & Verhoeven (2002, p.87).

Table 3. 2 Formant frequencies for Classical and Cairene Arabic (Newman & Verhoeven 2002, p.87)

Language Variety	i:		i		u:		u		a:		a	
	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
Classical	390	1725	440	1770	470	1120	480	1170	620	1455	616	1460
Cairene	290	1940	375	1575	290	830	360	912	610	1500	683	1435

The aim of the study was to determine the existence of High Classical style as an acoustically purer variety of MSA which could therefore be considered as a reference system and used as a tool for comparison when studying other varieties of Arabic. However, the results did not provide suitable evidence for that, since Classical Arabic vowels were not at acoustic extremes. Moreover, the difference in length between long and short vowels was found not to be statistically significant, contrary to what is found in MSA and other dialectal varieties. It was concluded, then, that High Classical Style vowels are not in a position to be used as a reference tool for other varieties of Arabic.

Vowels in other Arabic dialects, either individually or comparatively, have also received some attention from researchers. For instance, a study conducted by Alghamdi (1998) aimed at investigating whether the six vowels available in the standard variety of Arabic are realised in the same way when spoken by speakers of different dialects. The sample of the study included participants speaking Saudi, Sudanese and Egyptian Arabic. Alghamdi (1998) used 15 male informants who produced CVC syllables containing one of the six vowels found in Arabic. Syllables were produced in isolation because it was thought that they can be easily influenced by the context in carrier sentences. Syllables including long vowels were meaningful words while those containing short vowels were nonsense ones. However, no reason was given to the choice of nonsense words. All words were presented in a written form where short vowels were symbolised by diacritics.

With regard to the quality of the vowels, the study found a qualitative difference between long vowels and their short counterparts. Short vowels were more central than long ones, which tended to be more peripheral. More interestingly, a comparison of the results of the three groups showed that the MSA vowel system differs with regard to phonetic implementation from one dialect to another. Table 3.3 shows the mean formant frequencies for the three dialect speakers investigated by Alghamdi (1998).

Table 3. 3 Formant means of MSA vowels as produced in three Arabic dialects[5] (Alghamdi1998, p.8, 12, 16)

Vowel	i		iː		a		aː		u		uː	
Dialect	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
Saudi	402	1841	292	2286	573	1537	655	1587	451	1302	350	958
Sudanese	331	2066	272	2255	525	1564	635	1492	354	1308	319	984
Egyptian	357	1749	256	2175	468	1505	462	1677	370	1285	319	942

When the significance of these results was tested, some significant relations were found within and across dialects. For instance, the F1 in /u, uː, i/ was significantly higher for Saudi than for Sudanese and Egyptian. Moreover, the F1 for Egyptian for /aː/ was significantly lower than that for the Saudi and Sudanese counterparts. In the case of /a/ all dialects differed significantly with regard to F1. As for F2, it was significantly higher in Egyptian than in the two other dialects.

Within dialects, there was a significant difference in both F1 and F2 between short and long vowels except for F2 in Saudi and F1 of the Egyptian vocalic pair /a, aː/ (Alghamdi 1998, p. 6).

Another study that was concerned with Arabic vowels was conducted by Belkaid (1984), who investigated the realisations of these vowels in the Tunisian context. Belkaid collected data using 70 words containing a mixture of short and long vowels in CV and CVC contexts. The results of that investigation showed significant differences between short and long vowels. The average means of formant frequencies are shown in Table 3.4 below.

---

[5] All formant values have been rounded to the nearest full number.

Table 3. 4 Average means of vowels in Tunisian Arabic (Belkiad 1984, p.224)

Vowel	i		i:		a		a:		u		u:	
Formant values	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
	355	1830	285	2195	400	1640	425	1720	340	995	310	790

Another investigation of a group of Arabic dialects was conducted by Abou Haidar (1994) who compared eight Arabic dialects using one participant from each dialect. These dialects were Qatari, Lebanese, Saudi, Tunisian, Syrian, Sudanese, Emirati and Jordanian. Abou Haidar used CVC monosyllabic words. The average formants obtained from the study are presented in Table 3.5 below.

Table 3.5 Average means of vowels in some Arabic dialects (Abou Haidar 1994)

Vowel	i		i:		a		a:		u		u:	
Dialect	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
Tunisian	510	1690	315	2275	650	1590	**610**	1780	540	1135	360	830
Syrian	**415**	2135	330	2465	700	1680	710	1560	**430**	1200	320	**620**
Jordanian	565	1720	320	2295	780	1620	770	1521	580	1240	**260**	795
Saudi	540	1830	305	2530	695	1590	730	1540	540	1190	375	930
Sudanese	420	2000	325	2220	660	1600	730	1500	455	1040	380	900
Qatari	500	**1400**	**310**	**1990**	**620**	**1540**	621	**1280**	490	**1005**	310	830
Lebanese	490	1530	280	2010	640	1390	**610**	1430	475	1060	330	795
Emirati	460	1720	335	2065	640	1660	730	1380	475	1075	350	990

As can be observed from Table 3.5, there is a lot of variation in formant values across these dialects. However, the highest formant values (heavily shaded cells) are found in Jordanian while the lowest formant values (lightly shaded cells) are found in Qatari. Each dialect has at least one maximum or minimum value among the other dialects in one vowel or another, except for the Saudi dialect. The maximum difference in F1 between the vowels /i:-i/ is found in Tunisian (195 Hz) and the minimum difference is found in Syrian (85 Hz). The maximum difference in F2 in the same pair of vowels is found in Saudi (700 Hz) and the minimum difference is found in Sudanese (220 Hz). For the vowels /u:-u/, the maximum difference in F1 is found in Jordanian (320 Hz) and the minimum difference is found in Sudanese (75 Hz). As for F2 for the same pair of vowels, the maximum difference is found in Syrian (580 Hz) while the minimum difference is found in Emirati (85 Hz). Finally, the largest difference in F1 for the vowel pair /a:-a/ is found in Emirati (90 Hz) and the smallest difference is found in Qatari (1 Hz). For F2 for the same pair, the largest difference is found in Emirati (280 Hz) while the smallest difference is found in Lebanese (40 Hz).

One of the disadvantages of this study is the small number of participants. One participant from each dialect cannot be a true representative of that dialect. This also applies to other studies that have investigated Arabic dialects (e.g., Al-Ani 1970). Differences in results for the same dialect studied by two researchers can be attributed to this factor among others. Compare, for example, the results for the Saudi dialect obtained by Alghamdi (1998) to those obtained by Abou Haidar (1994) for the same dialect as shown in Table 3.6 below.

Table 3. 6 Different formant means for Saudi dialect obtained by different researchers.

Vowels	I		i:		a		a:		u		u:	
Formants	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
Alghamdi (1998)	402	1841	292	2286	573	1537	655	1587	451	1302	350	958
Abou Haidar (1994)	540	1830	305	2530	695	1590	730	1540	540	1190	375	930

3.1.1.1. Allophonic variation

The three main qualities of Arabic vowels, whether in their short or long forms, are phonetically realized in at least two allophones each (Al-Ani 1970, p.23); a plain as well as an emphatic allophone. Emphatic vowels are found in the environment of a group of consonants called the emphatics. These include /ðˤ, tˤ, sˤ, dˤ/, which have plain counterparts and are distinguished from them by having a secondary pharyngeal articulation simultaneous to their primary articulation. In the production of all the emphatics the tongue back is raised against the velum and retracts towards the pharyngeal wall (Abdul-Jaleel 1998, p.16). As a result of this, the emphatic vowels are more retracted and lower than their plain counterparts (Bukshaisha 1985, p.216).

In acoustic terms, the retraction of the tongue and the depression of its dorsum required by emphatics results in the lowering of the onset of the F2 (Card 1983, p.49) and the raising of F1 frequencies in the subsequent vowels because of the time needed for the tongue to move from a retracted position to the target vowel position (Hussain 1985, p.294). The emphasis effect is also anticipatory, so that a vowel can become emphatic even before the tongue reaches the position for the following emphatic consonant, as the examples given by Gairdner (1925, p.46) show.

(3.2) [ʃɑtˤ] (shore) [fɑsˤl] (division) [fɑːdˤ] (overflowed)

The same view was held much earlier by Goitein (1960, p.361) who reported that, in the Yemeni dialect, the effect of emphatic consonants goes beyond adjoining vowels and colours all vowels in the same sound unit. These findings have also been confirmed by other researchers such as Card (1983), Watson (1999), and Al-Masri & Jongman 2004). The results of the latter study showed that the effect of emphasis also spreads to neighbouring syllables both to the right and left of the target one although to a lesser degree. The finding that emphasis spreads in both directions was also confirmed by Card (1983) who also found that primary emphatics generally have lower F2 values than secondary ones. However, the spread of emphasis seems to be blocked when the high vowels /i/ and /u/ are found in the target syllable, i.e. the syllable with the emphatic consonant. Watson (1999) notes that the domain of emphasis spread varies from dialect to another in Arabic. For instance, while emphasis spread is limited to the adjacent vowel in Abha, a dialect spoken in Saudi Arabia, it spreads over the whole word in Qatari Arabic (Watson 1999, p. 290). However, the effect of emphasis spread was found to be limited to segments within the word and does not go beyond word boundaries (Card 1993, p. 60).

Another study that has dealt with emphasis spread in Arabic was conducted by Zawaydeh (1999) who studied the phenomenon in the context of the Jordanian dialect spoken in Amman. The aim was to investigate whether high sounds like /i/ and /u/ block spread in emphasis and, if so, in which direction. The results showed that leftward spreading is categorical while rightward spreading is gradient. Zawaydeh explained this difference by saying that leftward spreading is anticipatory and, therefore, it is stronger than rightward spreading which tends to be preservative. This is so since before articulating the emphatic, the speaker wants to make the contrast early by making the tongue back come closer to the uvula in order to produce the emphatic sound. This affects the articulatory configuration of the preceding vowel. However, the moment this emphatic consonant is produced, the speaker relaxes his articulators which lets F2 gradually rise towards a plainer vowel position (Zawaydeh 1999, p.174-75).

Spread of emphasis has also been studied by Davis (1993, 1995) in the Southern Palestinian context. It was found that emphasis spreads leftward to the end of the word while spread is blocked in the rightward direction by sounds like /i/, /ʕ/. The following are some examples given by Davis (1995, p. 474) where the underlining indicates places where emphasis has been blocked.

(3.3) /manaadiðˤ/ (ashtrays) /tˤ<u>iinak</u>/ (your mud) /ʕạtˤ<u>*ʕaan*</u>/ (thirsty)

The first word shows that emphasis spread to the left is not blocked in the whole word while that to the right is blocked by sounds like /i/ and /ʕ/. This was explained by Davis by saying that those sounds which have a +high feature are opaque to rightward emphasis spread. This spread leftwards also extends to prefixes if an emphatic is found in the stem as the following examples taken from Davis (1991, p. 4-5) for Cairene Arabic show.

(3.4) /hạ-sˤạllị/ (I will pray) /bạ-fạdˤdˤạl (I prefer)

The degree of the effect caused by emphatic consonants on vowels depends on the distance between the two articulatory positions of the two. Thus, front vowels are more affected than back ones because the distance over which the tongue moves is longer for /i/ than /u/, for example. Moreover, the results of Hussain's (1985) study also showed a small difference in F2 between /iː/ in plain and emphatic positions, while this difference was greater in the case of /i/. This was explained by the fact that /iː/ has a longer duration during which the tongue is capable of reaching its target position, which is not obtainable in the case of the short vowel /i/ in which the F2 never reaches its steady state and keeps falling throughout the vowel.

The degree to which vowels are affected by neighbouring emphatic consonants differs from one vowel to another depending on the quality of the vowel as well as its quantity. This is due to the fact that the influence of adjacent speech sounds depends on the difference in articulatory configuration between affected and affecting sounds; the greater the difference, the greater the effect (Rosner & Pickering 1994, p.22). For example, low vowels are more affected by neighbouring emphatic consonants than high vowels and, therefore, the qualitative difference between the emphatic and plain allophonic variations is greater in the case of [æ, ɑ] than those of [i, ị] and [u, ụ].[6] Thus, some Arabic native speakers cannot audibly make a distinction between the two elements of each of the two pairs [i, ị] and [u, ụ] while they can with the pair [æ, ɑ] (Muftah 2001, p.81). The presence of transition also depends on the position of the vowel relative to the back-front dimension. While the front vowel [iː] shows a transition, the back vowel [uː] lacks it (Ghazeli 1977, p.85). This reflects the degree of movement by the tongue from a pharyngealized consonant position to a vowel position. Moreover, when the transition is present, as in the cases of /iː/ and /i/, its length will have a positive relationship with the duration of the vowel; the longer the vowel, the longer the transition. Vowel duration also affects F2 frequency, with longer vowels having higher F2 (Ghazeli 1977, p. 85).

A more recent study that has dealt with emphasis in Arabic was the one conducted by Bin-Muqbil (2006) with the focus on MSA. The acoustic results of the

[6] A dot used under the symbol indicates that the vowel is emphatic.

study confirmed the common theory that emphatics exhibit greater coarticulatory influence on neighbouring vowels than non-emphatics. This is reflected in the reduction in the F2 of vowels in the vicinity of emphatics as a result of the dorsal retraction which marks the latter. Emphatic consonants were also found to cause high F1 transitions in neighbouring vowels. The study further showed that the emphatic consonants /ðˤ, tˤ, sˤ, dˤ/, unlike non emphatics, resist vowel-to-vowel coarticulation because the production of these four Arabic emphatics involves tongue backing which restricts movement of the tongue dorsum, which is the main articulator of vowels. This in turn hinders vowel-to-vowel coarticulation (Bin-Muqbil 2006, p. 223-25).

3.1.1.1.1. The status of the low vowel /a/

In addition to being found in the environment of emphatic consonants, the emphatic version of /a/ is also found adjacent to plain consonants like [b], [m] and [l](see section 3.2 below for examples in LA). Some researchers (e.g., Ferguson (1956) believe that these consonants are considered to be emphatics, which results in /a/ being emphatic when contiguous to them. However, the fact that the emphatic version of these consonants is only found in the environment of the low vowel /a/ may cast doubt on them being emphatic. A plausible alternative would be that the low vowel /a/ is emphatic and that these consonants are not.

Moreover, Ghazeli (1977) found no acoustic evidence in favour of the consonants [b], [m], and [l] being emphatic. He believes that it is the low back vowel /a/ which is pharyngealized and therefore its occurrence contiguous to consonants like [b], [m], and [l] will cause them to be backed (Ghazeli 1977, p. 134). However, if this vowel is pharyngealized, then why is its occurrence limited to the environment of these consonants?

On the other hand, Ghazeli believes that Arabic possesses only one low vowel which has two realisations: central in the eastern part of the Arab world including the Gulf region, Iraq and Jordan, and front in the more western countries including Egypt, Libya, Tunisia, Algeria and Morocco. In both cases this vowel is backed in a pharyngealized context. To explain the existence of the emphatic alongside the plain version of the low vowel alongside in these countries, Ghazeli maintains that its existence is limited to a small number of words most of which are borrowed and are not originally Arabic (Ghazeli (1977 p. 144). It is obvious from this review of Ghazeli's work that he makes a distinction between a pharyngealized low back vowel which may be found in the context of some consonants like [b, m, l] and a low vowel that becomes pharyngealized when occurring in the context of emphatic consonants like [tˤ, dˤ, ðˤ, sˤ]. The same view regarding the low vowel /ɑ/ being emphatic is held by Youssef (2006 p. 45) for Cairene Arabic.

Another consonant that is considered a secondary emphatic is the sound /r/ (Harrell 1957) which has been reported in a number of Arabic dialects. However, the /r/ being emphatic is conditioned by the context in which this sound occurs. For example this emphatic consonant loses its emphasis when preceded by the high long vowel /iː/. This alternation between emphatic and plain /r/ has been reported for the Cairene Arabic (Davis 1991; Younes 1993, 1994) and Tunisian (Ghazeli 1977) Arabic dialects. The loss of emphasis in /r/ when it is contiguous to a high front vowel can be explained by arguing that this sound is not fundamentally emphatic and that the emphatic version is only an allophonic variant of the sound. According to Ghazeli (1977, p. 166) the emphatic version of the Arabic /r/ is similar to the American

English retroflex/r/ as described by Delattre (1971) and which has a somewhat lower F2 and F3 than its plain counterpart as a result of tongue front depression and tongue back retraction.

The status of the emphatic version of the low vowel /a/ and that of the consonants [l], [m] and [r] in LA and other Arabic dialects is still in dispute and leaves the door open for more debate. Emphasis is not dealt with in this study due to space and time limitations. However, a study is currently being conducted (Kriba, in progress) which focuses on emphasis in LA.

3.1.2. Vowel length

Any of the three basic qualities of MSA vowels can have a long or a short form, with some corresponding difference in their quality as shown in section 3.1.1. The examples in 3.5 illustrate phonological length distinctions in Arabic (Muftah, 2001, p.83).

(3.5) Distribution of long and short vowels in Arabic

/i/	/zir/ button	/i:/	/zi:r/ large jar
/u/	/suq/ drive (imp.)	/u:/	/su:q/ market
/æ/	/fæhim/ he understood	/æ:/	/fæ:him/ knowing (adj.)

In Alghamdi's (1998) study of MSA vowels as produced in various Arabic dialects, the durations of the six simple vocalic phonemes available in MSA were measured. While there were major differences in the realisation of these vowels in terms of quality (see section 3.1.1), Alghamdi found that these vowels behaved in a similar way in all dialects with regards to quantity; the duration of long vowels was about twice that of their short counterparts. This big difference in duration between long and short vowels was attributed to the fact that the test syllables containing these vowels were produced in isolation.

Table 3. 7 Vowel duration ratios and means in Saudi, Sudanese and Egyptian dialects (Alghamdi, 1998, pp. 11, 15, 18, 19).[7]

Dialect	Vowel	i	i:	a	a:	u	u:	Mean
Saudi	Mean	111	248	133	311	114	137	
	Ratio	0.45		0.43		0.48		0.45
Sudanese	Mean	117	275	128	295	116	304	
	Ratio	0.42		0.44		0.38		0.41
Egyptian	Mean	98	255	122	316	110	253	
	Ratio	0.39		0.39		0.43		0.40

Table 3.7 shows the means and ratios of the duration of vowels produced in the three Arabic dialects studied by Alghamdi (1998). This considerable difference in duration between long and short vowels in Arabic was also found by Al-Ani (1970), who produced and measured vowels in isolation (Table 3.8). The duration ratio found by Al-Ani (1970, p.75) is similar to that found by Alghamdi (1998, p.6), Gairdner (1925, p.35) and Abumdas (1985, p.45), which suggests that Arabic long vowels are about twice as long as short vowels when produced in isolation or in short syllables.

Table 3. 8 Relative durations in ms. of Arabic vowels in isolation as spoken by Iraqi speakers (Al-Ani, 1970, p.23)

Vowel	i	i:	a	a:	u	u:
Duration	300	600	300	600	300	600

In a study by Mitleb (1984b), however, participants were asked to read a carrier sentence containing Jordanian monosyllabic minimal pairs with /i:, i/, /a:, a/

7 The formant mean values are rounded to the nearest full number.

and /u:, u/ vowel contrasts. The words used were /bi:t, bi:d/, /bitt, bidd/, /ba:t, ba:d/, batt, badd/, bu:t, bu:d/, /butt, budd/ and were presented in Arabic orthography. It was found that Arabic long vowels were only 35% longer than their short counterparts. Table 3.9 shows the average duration of Arabic vowels as spoken by Jordanian speakers (Mitleb 1984b, p.231).

Table 3. 9 Arabic vowel duration as spoken by Jordanian speakers (Mitleb 1984, p.231)

	Long vowels			Short vowels		
Vowel	/i:/	/a:/	/u:/	/i/	/a/	/u/
Duration	116	145	124	76	90	83
Mean	128			83		

According to the values in Table 3.9, the difference between long and short vowels is 45 ms and the ratio of short to long vowels is 0.65, which is higher than those found by Al-Ani (1970) and Alghamdi (1998) for the other Arabic dialects mentioned above, which were 0.50 and 0.42 respectively. This may be attributed to the methodology used in collecting the data. While Al-Ani collected data from isolated vowels and Alghamdi from monosyllabic words produced in isolation, Mitleb (1984b) used monosyllabic minimal pairs read in carrier sentences.

Vowel duration was also investigated by Hussain (1985) using two Gulf Arabic dialect speakers. The results of his study showed the existence of a prominent quantitative difference between short and long vowels. The ratio of the duration of short to long vowels was 0.55 in monosyllabic words. However, the mean difference between the short and long low vowels /a, a:/ was greater than that between the high vowels /i, i:/ and /u, u:/. Moreover, the degree of openness was also found to affect vowel duration. In this respect, the average duration of the short high vowels /i/ and /u/ was lower by 16% than that of the low short vowel /a/. The same difference was also found between the high long vowels /i:, u:/ and the low long vowel /a:/.

With regard to the effect of vowel height on duration, both Mitleb (1984b) and Hussain (1985) found a considerable difference in duration between low vowels and their high counterparts. As Table 3.9 shows, the low vowels /a, a:/ in Jordanian Arabic are longer than the corresponding high vowels by 17 milliseconds (14%), which was found to be statistically significant. The increase in duration from low vowels to high vowels with mid vowels having intermediate durations is consistent with the generally acknowledged linguistic tendency that relates vowel duration to the degree of vowel height (Lindblom 1967, Klatt 1976). A reasonable explanation for these differences is physiological. As Lindblom notes, "the temporal organization of speech sounds is determined by the amount of physiological energy that is consumed in producing them" (Lindblom 1967, p.22). That is, low vowels are longer than high vowels because they are produced by lowering the jaw which requires more energy.

3.1.2.1. The effect of linguistic context on duration

The effect of the context in which the vowel occurs on its duration has been dealt with by some researchers in terms of neighbouring sounds. In particular, the effect of voicing in the following consonant has been examined closely in many languages. Some universal tendencies have emerged and the duration of a vowel followed by a voiced consonant found to be significantly longer than that of a vowel followed by a voiceless consonant. However, in Arabic the results are not uniform.

While some studies (e.g. Port et al 1980, Hassan 1981, Alghamdi 1990) showed that this difference is statistically significant, others (e.g. Flege 1979, Hussain 1985, Mitleb 1984b) found a small but insignificant difference. These studies are reviewed below.

In Fleges's PhD dissertation (1979), the effect of consonant voicing on the preceding vowel was examined in Arabic using Saudi Arabic speakers. The results showed an insignificant effect. For example, the long vowel /a:/ preceding the two stops /d, g/ was only longer than that preceding /t, k/ by 6ms. This was attributed to the fact that Arabic possesses a phonemic length contrast for both vowels and consonants (Flege 1979, p.64) which should be maintained by speakers. Table 3.10 presents the mean duration of vowels before voiced and voiceless stops found by Flege for the Saudi Arabic dialect.

Table 3. 10 Mean duration of vowels before voiced and voiceless stops in Saudi Arabic dialect (Flege 1979, p.66).

	/t, d/		/k, g/	
	/ga:t/	/ga:d/	/ʕa:k/	/ʕa: g/
Mean	176.5	182.6	167.1	173.1
SD	24	28	17	22
Difference	6.1		6.0	

Regarding the effect of the preceding consonant on the following vowel duration, Flege found some evidence of temporal compensation between the vowel and the preceding consonant in word-initial position. It was found that vowels are significantly longer after voiced consonants than after voiceless ones. This was explained by the fact that voiceless consonants are longer than voiced ones, so in order to compensate for that, vowel duration is shortened after voiceless consonants. However, Flege admits that this effect needs further investigation (Flege 1979, p.67).

Mitleb (1984a) studied the voicing effect on vowel duration in Arabic. The aim of the study was to verify the widely recognized view that vowels are longer before voiced consonants than before voiceless ones. The study took three steps. First, relevant material from eight Jordanian Arabic speakers was elicited. This was in the form of CVVC words (kaas *vs* kaaz) produced in a carrier sentence (where VV is represented by the long vowel /a:/ and the final C is represented by the voiced or the voiceless consonants /z/ and /s/). A spectrographic analysis of these words showed no significant difference in vowel duration between those words ending with voiceless consonants and those ending in voiced ones. Then, the same procedure was conducted using comparable English words (cass *vs* kaz). elicited from two groups of eight native speakers of English and eight Jordanian speakers of English who had been in the United States for over three years at the time of the study.

Data produced by speakers of English showed that, unlike Arabic, English exhibits a difference in vowel duration depending on whether the following consonant was voiced or voiceless. Vowels preceding voiced consonants were found to be significantly longer than those preceding voiceless ones. In, Arabic on the other hand, no effect was shown of the voicing of the final consonant on the duration of the preceding vowel and, therefore, no significant durational difference between vowels preceding voiceless or voiced consonants was found.

Hussain (1985) studied the effect of the voicing of the consonant on the preceding vowel in Gulf Arabic dialect using two native speakers who produced /hVCC, hV:C/ minimal pairs of meaningful words . However, when meaningful words were not obtainable, nonsense words were used. It was found that the durational difference between vowels preceding voiced and voiceless consonants was not significant and amounted to only 10 ms on average. These results are in accordance with those found in the Jordanian dialect by Mitleb (1984a) and in the Saudi dialect by Flege and Port (1981).

While the above mentioned studies on Arabic found no significant difference in duration between vowels preceded by voiced consonants and those preceded by voiceless ones, others have found this difference to be significant. Alghamdi (1990), for example, investigated the effect of consonant voicing on both preceding and following vowel in the Ghamid dialect, an Arabic dialect spoken in Saudi Arabia. The results showed a significant difference in duration between vowels followed by voiced consonants and those followed by voiceless ones. Vowels preceding voiced sounds were significantly longer. Similar results were found by Port et al (1980) and Hassan (1981).

Alghamdi also found some effect of both the quality and quantity of vowels on their duration when occurring in the environment of voiced and voiceless consonants. For instance, long vowels showed a slightly bigger proportional difference than short vowels. That is, the difference between long vowels followed by voiced consonants and those followed by voiceless ones is bigger than the difference between short vowels followed by voiced consonants and those followed by voiceless ones. The same thing was found between low and high vowels; low vowels were longer than high ones. For example, /a:/ shows a bigger difference than /i:/.

Regarding the consonantal voicing effect on the duration of a following vowel, the results showed that vowel duration is significantly shorter after voiceless stops than after voiced ones. The ratio of the vowel preceding voiceless consonants to that following voiced ones in monosyllabic words was 0.85.

The influence of consonant place and manner of articulation on vowel duration was also examined by Hussain (1985). Regarding the place of articulation of the following consonant, there is a general tendency for vowel duration to increase with the transition time required for the articulator to move from the vowel position to the consonant position. That is, vowel duration is constrained by the distance between the point of articulation of the vowel and that of the following consonant. For example, it was found that the front vowels /i, i:/ preceding voiced stop consonants take the following durational ranking /g/ > /d/ > /b/. These vowels being longer before /g/ than /d/ is due to the fact that the distance over which the tongue moves from the vowel position to the velar position for /g/ is longer than that to the dental position of /d/. On the other hand, these vowels are shorter when preceding /b/ than when preceding /d/ because the tongue is not involved in the production of the bilabial stop /b/ and, thus, no tongue movement is required. However, some tongue movement is needed when these vowels are followed by the dental stop /d/. Manner of articulation, on the other hand, did not show any consistent pattern regarding its effect on vowel duration (Hussain 1985, p.151).

To sum up, the literature review of the effect of consonantal voicing on duration in Arabic shows that this effect is still debated among researchers. This controversy might be attributed to several factors including the methodology followed by researchers in collecting data, such as the number of participants and the material used. Another factor that might have played a role is the different varieties of Arabic that have been studied and which may exhibit different dialect-specific characteristics.

3.1.3. Diphthongs

The MSA sound system contains only two diphthongs: a fronting diphthong /ai/ as in /kaif/ *how*, /ʕain/ *eye*, /bait/ *home* and a backing diphthong /au/ as in /haul/ *terribleness*, /maut/ *death*, /laun/ *colour* (Muftah 2001, p.83). The starting positions of the two diphthongs are the same, which is the low vowel /a/, while they differ only in their second element which is the high front vowel for /ai/ and the high back vowel for /au/. The second element of these diphthongs has sometimes been transcribed using the semi-vocalic elements /j/ and /w/. For example, Gairdner (1925, p.45) used these semi-vocalic notations to represent the second element of the diphthongs, but he limits this use to Classical Arabic since, according to the researcher, the glide goes up to a consonantal position in the production of these diphthongs in this Arabic variety.

On the other hand, Dickens (1996, p.7) considers [i] and [j] as variants of the same vocalic phoneme. The same thing applies to [u] and [w]. According to him, these allophones do not contrast with one another in the same context and, thus, the two transcriptions [naum], [nawm] or [bait], [bajt] refer to the same word. This view is also supported by Roca and Johnson (1999), who state that [w] and [j] are vowels that differ only in their length from /u/ and /i/ as far as articulation is concerned. That is, /w/ and /j/ are short vocalic variants of /u/ and /i/ (Roca & Johnson 1999, p.269). In spite of their phonetic articulation, which is similar to vowels, these two sounds have a phonological distribution similar to that of consonants.

3.1.4. The effect of speaker variability

Al-Tamimi & Barkat-Defradas (2002) conducted a study investigating intra-speaker and inter-speaker variability in speech production and perception in Jordanian and Moroccan Arabic. The aims of the study were to understand the relations between the production and perception of vowels, to study differences in the production and perception of the vowels in male and female speakers and listeners, and finally to study the difference between long and short vowels in these Arabic dialects. To the best of the present researcher's knowledge, this is the only study that has dealt with perception as well as production in Arabic. 35 male and female speakers each of these dialects participated in the study in which 8 vowels for Jordanian Arabic (i, i:, u, u:, e:, o:, a, a:) and 5 vowels for Moroccan Arabic (i:, u, u:, ə, ā were involved.

In the production task, vowels were presented in a CV structure in 51 items for Jordanian and 35 for Moroccan. However, the researchers did not clarify whether this structure used meaningful words or nonsense ones. 10 male native speakers of each variety were recorded producing these items. On the other hand, the perception task was based on a method of formant adjustment of isolated synthetic vowels as used by Johnson et al (1993). In perception, participants were presented with 160 tokens for Jordanian and 100 tokens for Moroccan and their task was to find the best prototypes that optimally represented the vowels in each dialect. The production data was analysed by measuring F1 and F2 from the middle portion of the vowels. In perception, vowel formant values were converted from Hertz to Bark before calculating the average means of these formants.

The results showed that the same pattern of realisation was found in production and perception in that short vowels tended to be central while long ones tended to be peripheral. This acoustic difference between long and short vowels was found to be statistically significant. However, greater variability was observed in Moroccan, which has fewer vowels than Jordanian. The perception experiment confirmed previous findings by Johnson et al (1993) that, with regard to F1, closed vowels are more closed in perception while open vowels are more open. The same thing applies to F2. That is, back vowels are more back and front vowels are more front, which means that perceptual vocalic space corresponds to a hyper-articulated space. That is, the perception area is larger than the production area in order to integrate speech variation.

The discussion of the literature related to Arabic dialects reveals some discrepancies when the results of different studies are compared. These discrepancies range from differences in vowel formant values to differences in durations which might result in different realisations of vowels from one dialect to another. Phonological differences such as those related to the effect of context are also found. Several factors are involved. First, the wide area occupied by the Arab world where these dialects are spoken is huge which means that the dialects sometimes exhibit big differences among them. Second, MSA is often shaped by the native dialect of its speakers which makes it difficult to arrive at accurate specification of the characteristics of the variety as a whole. It is often the case that MSA spoken by Iraqis, for example, sounds different from that spoken by Egyptians. This in turn makes generalisations derived from cross-linguistic comparisons hard to achieve. For example, a study comparing Saudi Arabic with English might give different results from that comparing English with Moroccan under the same circumstances. Finally, the different methods and procedures adopted by different researchers not only complicates cross-linguistic comparisons but also results in problems when comparisons between Arabic dialects are conducted.

3.2. The Libyan Arabic vowel system

In addition to the six vowels that are found in MSA and also in LA (see section 3.1.1), there are other vowels which are only found in the Libyan dialect and other Arabic dialects. These are the mid front long vowel /e:/ and the mid back long vowel /o:/ (Abumdas 1985, p.43). The words in (3.6) are some examples of the vowels that form part of the Libyan vocalic system.

(3.6)	*vowel*	*word*	*gloss.*
	/i:/	/gi:s/	measure
	/i/	/mis/	touch
	/u:/	/mu:s/	knife
	/u/	/bun/	coffee
	/e:/	/be:t/	home
	/o:/	/mo:t/	death
	/a:/	/fa:t/	passed
	/a/	/ran/	rang

The short vowel [o] is also found in dialects spoken in the north east of Libya. The following words are some examples from Aurayieth (1982, p.23).

(3.7) [ʔilbiso] they dressed, [ʔimsiko] they held [ʔigsimo] they shared.

However, what is thought by Aurayieth to be an independent vowel is merely an allophone of the long mid back vowel /o:/. This short allophone is only found in the final positions of words, and it becomes long elsewhere. For example, when the words in (3.7) are followed by the object pronoun /ha/ (*it),* they are realised as the ones in (3.8) below.

(3.8) /ʔilbisoːha/ *they wore it*, /ʔimsikoːha/ *they held it*, /ʔigsimoːha/ *they shared it*

As for vowel length, all vowels found in LA show length contrast except for the long mid front vowel /eː/ and the long mid back vowel /oː/. The corresponding vowels for these in MSA are originally diphthongs (Gairdner 1925, p.42) but many modern-day dialects have monophthongal realisations instead as the examples in (3.9) show. The dialectal words in the left column are produced in MSA as those in the right column.

(3.9)	LA	MSA	gloss.
	/beːn/	/bain/	between
	/ʕeːn/	/ʕain/	eye
	/loːn/	/laun/	colour
	/koːn/	/kaun/	universe

Botagga (1991) investigated LA vowels as spoken in Sebha, the third major city in Libya located in the south of the country (see map in Figure 1.1 in Chapter One). He also adds /ʌ/ as an independent phoneme to the LA vowel inventory and excludes /oː/ from it (Botagga 1991, pp.70-73). Below are examples that were given of the /ʌ/ vowel.

(3.10) /bʌṭṭɑ/ duck /gʌṣr/ palace / ṣʌbr/ patience

However, a close inspection of these words reveals that they all contain emphatic consonants, which may have affected the neighbouring vowel and contributed to its realisation as [ʌ]. It is clear from the examples above that the vowel [ʌ], if found in Libyan Arabic spoken in Sebha, is merely an allophone of the low central vowel /ɑ/. In fact, Botagga contradicts himself, saying that "/æ/ and /ʌ/ are two variants of the same phoneme /ɑ/" (Bottaga 1991, p.70) in another part of his thesis.

Abumdas (1985, p.47) also considers the emphatic forms of LA vowels as independent phonemes and not merely allophones of these vowels. Minimal pairs are given as examples for the plain and emphatic forms of /ɑ/ to demonstrate that they are independent phonemes and not allophones belonging to the same phoneme, as can be seen in (3.11) below.

(3.11)	*Plain*	*Gloss.*	*Emphatic*	*Gloss.*
	/bællæh/	he wet	/bɑllɑh/	by God
	/wællæ/	he returned	/wɑllɑh/	by God
	/næbbæh/	he called sb. attention	/nɑbbɑh/	flirting
	/bæ:bæh/	his door	/bɑ:bɑh/	father
	/bæ : læh/	his attention	/bɑ : lɑh/	bundle/ bale
	/jæ : ri/	running	/jɑ : ri/	my neighbour
	/bæħħæh/	hoarseness	/bɑħħɑh/	it is finished (baby talk)

However, apart from the last example in which both the plain and the emphatic forms of the vowel occur in the environment of the guttural sound /ħ/, all the other sounds are not found in the environment specified by Abumdas for emphatic vowels, which is in the vicinity of emphatic and guttural consonants. This might appear to support the view that these vowels are independent phonemes and their acoustic status is not due to their occurrence in the environment of emphatic consonants. The examples in (3.11) might superficially indicate that both plain and emphatic versions of /ɑ/ can be considered separate phonemes regardless of the environment in which each occurs, and that each version also has a long and a short form which are also regarded as separate phonemic elements.

The question that arises now is that, since the emphatic version of the low vowel is considered an independent phoneme because of its occurrence contiguous to plain consonants, there would be no reason for the opposite situation not to take place, i.e., the occurrence of the plain version of the vowel in the emphatic consonantal environment. This, however, is not the case in LA. The plain low central vowel is only found in the plain context in LA. It can therefore be concluded that the plain and emphatic versions of the low vowel /a/ are not independent phonemes but only allophones of the same vowel. This is supported by the fact that the emphatic low vowel /ɑ/ is only found in the vicinity of plain consonants in a very small number of words.

Moreover, the status of some of what are thought to be plain emphatic consonants is often debatable. For example, the two consonants /l/ and /b/ found in the words in (3.11) are considered emphatics by some researchers into some Arabic dialects such as Lebanese (Obrecht 1968, Nasr 1959) and Palestinian (Card 1983).
In terms of emphasis, the distribution of emphatic vowels in LA is the same as that of MSA emphatic vowels; they are found in the environment of emphatic consonants (see, for example, Aurayieth 1982, p.24). However, with regard to vowels contiguous

to guttural consonants, Libyans vary in their production. For example, the low vowel is produced [æ:] in some dialects while it is produced [ɑ:] in others in the same context and with the same meaning. According to the knowledge of the present researcher, the following words are found across Libyan local dialects.

(3.12)	*plain*	*emphatic*	*gloss.*
	[gæ:l]	[gɑ:l]	said
	[ɣæ:li/]	[ɣɑ:li]	expensive
	[xæ:nib]	[xɑ:nib]	thief

However, these two vocalic varieties may lead to different meanings when used in the same context. The following words are examples given by Abumdas (1985, p.47) from a dialect spoken in the city of Zliten, 150 km to the east of Tripoli.

(3.13)		
	/xæ:li/ empty	/xɑ:li/ my uncle
	/kæ:ri/ tenant	/kɑ:ri/ my value
	/gæ:l/ he exempted	/gɑ:l/ he said

The emphatic vowel in these examples also affects the pronunciation of the preceding and following consonants. Thus, these consonants in the environment of this vowel are somewhat retracted as a result of emphasis spread from the neighbouring vowel. As a result, these consonants may sound different when contiguous to plain vowels from when they are adjacent to emphatic ones. For example, the /l/ in /xæ:li/ sounds clear while that in /xɑ:li/ sounds dark.

As has already been stated, plain and emphatic forms of the low vowel /a/ are not independent phonemes but allophones of the same phoneme which are distributed complementarily in spite of the small number of words in which the emphatic version of the vowel occurs in the plain context.

Moreover, the examples in (3.12 and 3.13) cast some uncertainty on the relations of emphasis between vowels and consonants. While some studies claim that vowels are the source of emphasis effects, others say that it is the emphatic consonants which affect neighbouring vowels (Aurayieth 1982, p.24). The question that arises is that, if emphatic vowels are so because of neighbouring emphatic consonants, then why are these emphatic vowels sometimes found in non-emphatic contexts? Moreover, a distinction should be made between vowels being emphatic and vowels being back, in that back vowels are not necessarily emphatic.

It is worth noting here that these descriptions of Libyan vowels are based mainly on auditory studies, some of which have been mentioned above. One aim of this study is to analyze these vowels acoustically and arrive at more comprehensive descriptions of them which take into consideration their acoustic properties in addition to their articulatory and auditory features.

3.2.1. Vowel distribution

In addition to emphasis as a factor affecting vowel distribution, there is another important factor involved in the distribution of vowels. This factor is the position occupied by the vowel in the word. For example, LA vowels do not occupy initial positions in a word. In fact, vowels in initial positions of words are also not found in MSA. However, while the occurrence of consonant clusters word-initially is not permissible in MSA, it can be found in LA and other Arabic dialects. When two consonants occur initially in a word in MSA, a default vowel is inserted before the first consonant to break that cluster. And since the occurrence of vowels word-initially is disallowed, the default consonant /ʔ/ is inserted before the default vowel as in the examples in (3.14) taken from Muftah (2001, p.134).

(3.14) *Underlying form*		*Surface form*	
/stagbal/ →	/istagbal/ →	/ʔistagbal/	to receive
/ndaθar/	/indaθar/	/ʔindaθar/	to extinguish
/ltamas/	/iltamas/	/ʔiltamas/	to appeal

However, word-initial consonant clusters are found in LA, as some of the examples in (3.15) taken from Aurayieth (1982, p. 26) show. These examples in (3.15) also show the vowel distribution in the words in LA.

(3.15) vowel distribution in words in LA

Position	*initial*	*medial*	*final*
Examples	-	/mæ:l/ (money)	/mʃe:/ (he went)
	-	/dib/ (walk slowly)	/di:b/ (wolf)
	-	/sxu:n/(hot)	/kbur/ (he grew)

LA vowels have more distributional variations than those mentioned so far. These variations are determined by environmental factors. For example, the vowels /æ/ and /ɑ/ may be omitted when unstressed in some dialects such as TA or realized as [ə] or [ʌ] in other dialects such as DA. Compare the words found in TA to those found in DA in (2.16).

(3.16) TA+DA	Gloss	TA	DA	Gloss
[fækkær]	he thought	[fækkru]	[fækkəro]	they thought
[fɑdˤdˤɑl]	he preferred	[fɑdˤdˤlu]	[fɑdˤdˤəlo]	they preferred

As the examples above show, the omission of the vowel in TA or its reduction to [ə] in DA is caused by the unstressed context of the vowel when the possessive pronoun (*they*) is added finally to the word. It should also be noticed that the vowel representing this possessive pronoun is realized differently in the two dialects. While it is realised as [u] in TA, its articulation is [o] in DA. This short vowel in the final position of the word is originally long when it is in medial positions. When the two words [fɑdˤdˤlu] and [fɑdˤdˤəlo] are followed by the object pronoun it, they are pronounced together with this pronoun as [fɑdˤdˤlu:h] and [fɑdˤdˤəlɔ:h] *they preferred it)* respectively. As mentioned above, the short vowel /o/ is an allophonic variant of the long vowel /o:/. This long vowel turns short when it occurs in the final positions of words.

While emphasis and position of the vowel in the word are determinants of vowel distribution, while vowel type is a determinant of word stress position. In fact, stress is attracted by heavy syllables and disfavours light ones. Because long vowels and diphthongs affect syllable weight in the sense that syllables containing long vowels are considered heavy, these syllables are often seen as candidates for being stressed. This is further illustrated by the examples in (3.17) taken from Aurayieth (1982, p.23).

(3.17) /mú : s/ (knife) /musé : n/ (two knifes)

/rá : kib/ (he is riding) /rakíbha (he is riding it)

/saiyá : ra/ (car) /saiyará : t/ (cars)

Stress in LA as postulated by Al-Ageli (1995, p.212) goes from right to left in its search for a suitable domain in the word and, therefore, when there are two long vowels in the word it is the one on the right which is stressed as in /saiyá : ra/ (car) and /saiyara : t/ (cars). The movement of stress towards the end of the word is available across languages and motivated by a language-specific rule known as end stress (Roca and Johnson 1999, p.322). However, when no long vowels in the word are found stress will be attracted by closed (i.e. heavy) syllables (CVC) rather than open (i.e. light) ones (CV) as in /rakíbha (*he is riding it*) which is syllabified as CV.CVC.CV.

3.2.2. LA diphthongs

As said before (see section 3.0 Chapter Three), a distinction is often made between MSA the standard variety of Arabic officially and formally used in all countries constituting the Arabic world and Arabic dialectal varieties which are mainly informally and colloquially used. The focus here is on the regional type of dialects since "[r]egional variation in the way language is spoken is likely to be one of the most noticeable ways in which we observe variety in language" (Wardhaugh 2002, p.43). These dialects differ from one country to another in the Arab world. For example, the dialect spoken in Egypt sounds differently from that spoken in Libya and the latter also sounds differently from the one spoken in Syria. The differences between these regional dialects are not restricted to differences in aspects of pronunciation but also cover features of vocabulary and grammar. It should be emphasised, however, that the regional and political boundaries between these countries do not always coincide with boundaries between these dialects. That is, there is no one particular point at which the changeover from one dialect to another takes place. Rather, areas around these political and regional boundaries constitute transitional areas from one dialect to another forming a situation that is often known as dialectal continuum (Wardhaugh 2002, p. 44). Therefore, residents on the two sides of each boundary often have fewer differences in speech from others who reside at a further distance from the same boundary. For features charactering Arabic dialects, this study will mainly rely on findings obtained by other studies conducted by other researchers in spite of the fact that these features might not be found in the whole region which represents a certain dialect or another. Finally, it should also be emphasised that whenever the term 'Arabic dialects' is mentioned, the focus is on the regional dialects and not on any other type of dialects.

Regarding diphthongs, in addition to the two MSA diphthongs mentioned above (/ai/ and /au/; see section 3.1.3) which are also found in the Libyan dialect, there are others which are found in Arabic dialects including LA, but not in the standard variety of Arabic known as MSA. Abumdas (1985, p.83) believes that there are four other diphthongs that are available in Libyan Arabic dialects. These are /uj, uw, ij, ew/. The /j/ and /w/ symbols used by Abumdas in these diphthongs are considered by Dickens (1996, p.7) as consonantal variants or allophones of the vocalic phonemes /i/ and /u/ respectively. According to Gairdner (1925, p.45), these consonantal allophonic variations are only found in classical Arabic. However, as far as articulation is concerned, /w/ and /j/ differ only in their length from /u/ and /i/. That is, /w/ and /j/ can be considered as short vocalic variants of /u/ and /i/ (Roca and

Johnson 1999, p.269). In spite of their phonetic articulation which is similar to vowels, these two sounds have phonological distributions similar to those of consonants in that they occupy the margins of the syllable rather than the nucleus, which is often occupied by vowels.

However, when the two diphthongs argued to be found in LA by Abumdas were auditorily compared to their counterparts in MSA, it was found that the first element in the two diphthongs in LA is shorter than the second element, which is opposite to the situation in MSA. This is obvious, for example, when the two words /jai/ (coming) and /rai/ (opinion) found in the dialect are compared to /jaib/ (pocket) and /raib/ (doubt) in MSA. It can also be observed that the second element in MSA is tenser than the first element of each diphthong. This might justify the use of the /w/ and /j/ as symbols for these elements in MSA but not in LA. It should be emphasised however that these auditory observations need to be enhanced with articulatory and acoustic investigation to assess their real status, which is beyond the scope of this study. Abumdas (1985) provides the following examples of the four diphthongs mentioned above found in LA

(3.18)	*Diphthongs*	*Examples*
	/uj/	/xuj/ my brother
	/ij/	/rijtak/ your lung
	/ew/	/xtewa/ little step
	/uw/	/daluwkum/ your bucket (pl.)

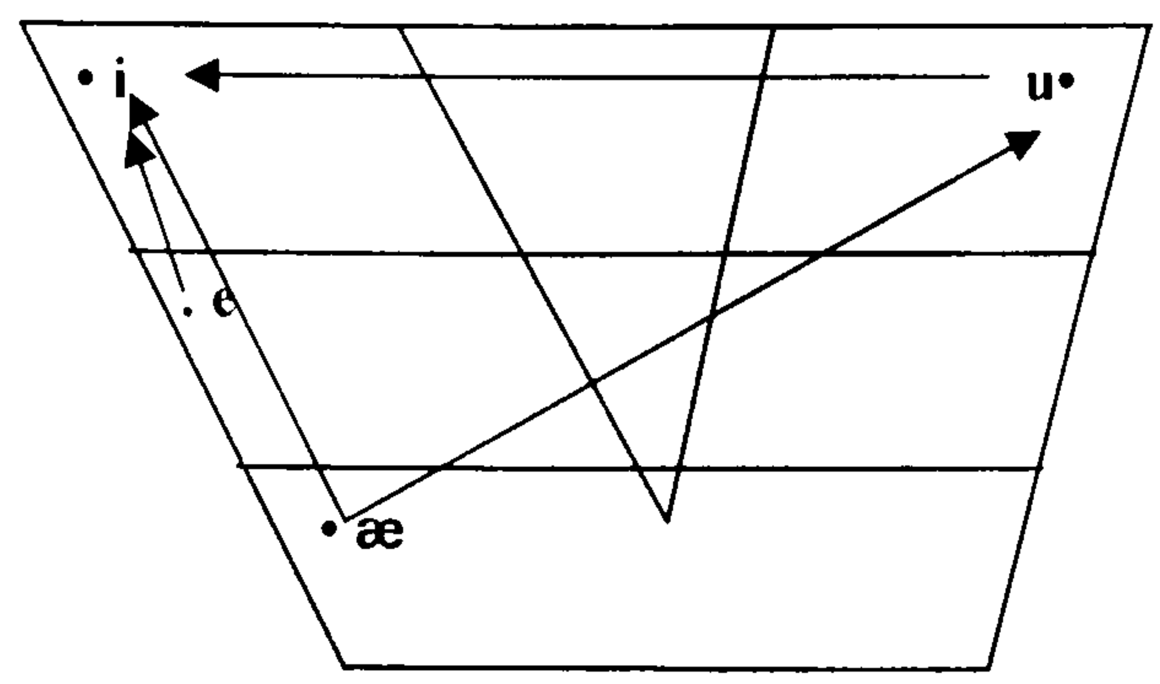
The symbols used in these examples for these diphthongs are the same ones used by Abumdas who did not clarify in what sense /ij/ and /uw/ are diphthongs. Based on the discussion above which considers the two elements of each of the pairs /i/-/j/ and /u/-/w/ as phonetically the same, the two vowels /ij/ and /uw/, if found in the dialect, should rather be considered long simple vowels rather than diphthongs and, therefore, could be symbolised instead as /ii/ and /uu/ respectively. Moreover, a close examination of the words above reveals that only the first two, i.e. /xuj/ and /rijtak/, contain diphthongs. The other two, /daluwkum/ and /xtewa/, do not include any complex vocalic elements. The realization of the word /daluwkum/ in MSA is [dalwukum] and it is syllabified as /dal.wu.kum/. This word is realized in the Libyan dialect as /dalukum/ after applying a dialect-specific deletion rule, as can be shown in (3.19) below:

(3.19)	MSA		LA
	/dalwukum/	$\longrightarrow$	/dalukum/

The word /xtewa/ is only found in the dialect and can be syllabified as /xte.wa/. This syllabification of the word shows that the word does not contain a diphthong. The sound /w/ is not the second element of an /ew/ diphthong but a semivowel occupying the onset of the second syllable of the word. Regarding the complex vowels in /xuj/ and /rijtak/, my intuition as a native speaker is that they are better phonetically transcribed as [xui] and [reitak].

Based on this discussion, Libyan diphthongs are four in number. Three of these are fronting as in /neitak/ (your intention), /ræi/ (opinion), /xui/ (my brother), and one is backing as in /kæu/ (cable). These diphthongs can be represented in the vowel quadrilateral in Figure 3.3 below.

Fig.3. 3 LA diphthongs



It is worth noting, however, that the representation of the diphthongs is schematic and an approximation, not based on any systematic analysis. .

The occurrence of the complex vocalic segment /ui/ is very rare in the dialect. A search for words containing this vowel yielded only two examples: /bui/ *my father* and /xui/ *my brother*. Moreover, the vowel is not available in MSA where these two dialectal words /bui/ and /xui/ are realized as /ʔæbi/ and /ʔæxi/ (the final /i/ in both words derives from the morphology and represents a possessive pronoun meaning *my)*.

3.3. Previous Libyan Arabic studies and their limitations

A review of studies related to Libyan Arabic phonology reveals that they are small in number and were mainly conducted in the 1980s. In his dissertation entitled "*The Phonology of the Verb in Libyan Arabic*", Aurayieth (1982) discussed DA. As can be noted from the title, the verb is the main concern of the study. However, vowels were discussed in some detail in a separate section in the thesis. The analysis of vowels and their distribution was conducted within the framework of generative phonology introduced by Chomsky and Halle (1968), focussing mainly on their classification according to their distinctive features. First, a list of the LA vowels was presented with descriptions and examples. Then, general observations about these vowels and their distribution were given.

Abumdas's (1985) dissertation entitled "*Libyan Arabic Phonology*" discussed the Libyan Arabic variety spoken in the city of Zliten located to the east of Tripoli, the capital city. A detailed impressionistic description of the sound system of the dialect was given. Distributions and classifications of the vowels into simple and complex, short and long, and plain and emphatic were also provided.

Another relevant study was conducted by Elgadi (1986). The work focused on the Tripolitanian dialect and classified its sounds using a generative approach similar to that followed by Aurayieth. Most of the analysis was conducted within the framework proposed by Chomsky and Halle (1968) and, thus, the sounds of LA in Elgadi's work were described in terms of their distinctive features and the rules governing the changes these sounds undergo. However, the main focus was on a suprasegmental phonology of the dialect and, therefore, phonological processes such as gemination, metathesis and assimilation were discussed. Syllable structure and stress assignment rules in the Tripolitanian dialect were also dealt with.

Another study which emphasised the suprasegmental features of the Tripolitanian dialect was conducted by Al-Ageli (1995). The work was entitled "*Syllabic and Metrical Structure in Tripolitanian Arabic*" and took the form of a comparative study in the framework of the Standard and Optimality theories. The Standard Theory is the one based on Chomsky and Halle's (1968) groundwork 'the Sound Patterns of English' under which sounds are contrasted according to their distinctive features. The study aimed to provide an exhaustive analysis of stress and syllable structure in TA, and to test the propositions of optimality theory in the context of the dialect (Al-Ageli, 1995, p.viii). However, only two chapters from the six in the thesis were devoted to TA. Chapter four was concerned with syllables in TA while chapter six dealt with stress in the dialect. The remainder of the thesis was mainly a review of phonological theories including the Standard Theory, Syllable Theory, Stress and Metrical Theory and Optimality Theory.

Finally, Laradi's (1983) work, entitled "*Pharyngealisation in Libyan (Tripoli) Arabic*", focused on the physiological and articulatory aspects of the phenomena of pharyngealization in TA, the Arabic dialect spoken in Tripoli. The study mainly relied on observations made on video-endoscopic and video-fluorographic recordings of words and phrases containing uvular, pharyngeal and pharyngealized consonants in initial, medial and final word positions. The main finding of this study was that epiglottal and pharyngeal constrictions are the main articulatory features required for the production of pharyngeal sounds in Libyan Arabic. With regard to vowels it was concluded that "pharyngealized consonants have the effect of lowering the close vowels and retracting the open vowels whether they precede or follow the other consonants" (Laradi 1983, p.23).

To conclude, it is obvious that research conducted on LA phonology has been very limited. Moreover, the majority of existing studies have employed theoretical and impressionistic approaches in deriving the properties of the language. When experimental methods were used, as in Laradi's (1983) work, the focus was not on vowels. It is obvious from our review of these works, therefore, that LA vowels have not so far been exclusively covered in any research. The present study extends knowledge on LA and deals specifically with LA vowels. It is the aim of this study to investigate LA vowels in more detail using experimental techniques in order to present a comprehensive account of their realisation and to explore the acoustic properties that are relevant to their production and perception.

3.4. Summary

This chapter reviewed the literature related to the Arabic vocalic system. It consisted of three main sections. Firstly, the Arabic vowel inventory was dealt with. Vowel quality and quantity were considered in some detail. Studies related to the effect of context in terms of neighbouring consonants on vowel duration and the effect of speaker variability were also reviewed. Secondly, the LA vowel system was described. In this respect, vowel distribution and the diphthongs found in LA were considered. The final section was devoted to reviewing previous Libyan studies and their limitations. In the following two chapters, the methods used in the present research are presented and a rationale for the design of the instruments used to collect both production and perception data is given. Then the results are presented and discussed.

PART TWO

METHODOLOGY, RESULTS AND DISCUSSION

CHAPTER IV
PRODUCTION TASK

CHAPTER V
PERCEPTION EXPERIMENT

CHAPTER VI
GENERAL DISCUSSION AND CONCLUSIONS

CHAPTER IV

PRODUCTION TASK

4.0. Introduction

As mentioned in the Introduction Chapter this study aims to investigate the Libyan Arabic vowel system as acquired by Libyan Arabic native speakers. Therefore, this research is concerned with studying the vowels contained in this system and exploring their phonetic properties. In this respect, an acoustic investigation of the quality of these vowels as well as their quantity is conducted. The acoustic analysis is supported by an auditory examination of the vowels. The combination of acoustic and auditory analyses is motivated by the fact that neither of these types of analysis is capable on its own of giving a comprehensive picture of the status of these vowels. Most of the studies conducted so far regarding LA have been auditory in nature. Those which have followed acoustic procedures were mainly concerned with consonants rather than vowels (see section 3.3 in Chapter Three).

Another issue that has never been studied is how these vowels are perceived. This study therefore constitutes the first step in examining the relationship between vowel perception (Chapter Five) and vowel production in further research on the vowel system of the dialect.

4.1. Design of the production task

The methodology used here is a combination of quantitative and qualitative techniques. In order to answer the first part of the main research question "What is the vowel inventory of LA?" detailed research was conducted in the literature (see Chapter Three). That research has shown that there are eight Libyan Arabic vowels.

However, considerable disagreement over the properties of these vowels was noted. This is partially due to the nature of the studies conducted so far which have mainly relied on the use of impressionistic auditory analyses. Another factor is the inter-dialect variation in producing these vowels. Libya occupies a vast area and, therefore, has a large number of dialects in spite of the small size of a population of less than 6 millions.[8]

In order to extend previous impressionistic work, this study adopted an experimental technique, which allowed greater objectivity and reliability. While acoustic analysis has been used widely in phonetic research, it has never been adopted for the study of LA vowels so far.

In dealing with the inter-dialectal problem, the study focuses on one dialect in order to present a comprehensive account of the vowels in this dialect with a robust sample (see sections 4.1.2 & 5.1.3) that would allow the later study of dialectal variation. In both production and perception tasks, a relatively large number of participants were chosen in order to make it possible to generalise the findings. Moreover, in this study, auditory analysis is employed simultaneously with acoustic analysis. Vowels are best described in terms of their acoustic properties rather than in terms of their articulation (Ladefoged 2003, p.104). However, although acoustic analysis tends to be more objective and reliable than auditory analysis (Fant 1973, p.3), acoustic analysis may have its own difficulties when it comes to describing vowels. Another problem concerns the fact that different acoustic measurements for the same data can be obtained depending on the programme used, the setting, and the recording equipment (see, for example, Harrison 2004). Differences in vowel formant measurements can also be due to the size of the vocal tract among the speakers

---

8 Libya 2003 census of population. http://en.ljbc.net/online/lypop.htm. [Accessed on 27.2.2008].

participating in an acoustic study. For instance, an adult often has lower formant frequencies than a child and a woman has higher formants than a man (Fant 1973, p.84). However, some effort has been devoted by researchers to accounting for these differences by performing *normalisation* on acoustic measurements (Watt & Fabricius 2003, p.2). This is a process that minimises these differences to a certain extent. (For more information on normalisation, see section 2.7.3 in Chapter Two.)

4.1.1. The language variety under investigation

As stated before, Libya occupies a very vast area (1,759,540 sq km).[9] Therefore, considerable dialectal variation is found throughout the country. However, these dialects can be grouped into three major dialects (see Section 1.1 in Chapter One). One of these major dialects is that spoken in the Tripolitania region in the north west of the country which also includes the capital city of Tripoli. In order to control the variable related to dialectal variation, this study focuses on the dialect spoken in Rayaina, a city found in the region of Tripolitania. This is due to it being the author's dialect so that the investigation would benefit from a native-speaker's intuition. Moreover, the dialect bears similarities with other dialects found in the same region (i.e., Tripolitania) which makes it possible for the results of this study to be representative of features of the wider region and therefore of interest to a wider audience.

The city of Rayaina is located 142 km to the south west of the capital city of Tripoli (see map in Figure 4.1)[10]. It is a small city with a population of less than 50,000 people located on Nafusah Mountain, one of the highest mountains in Libya.

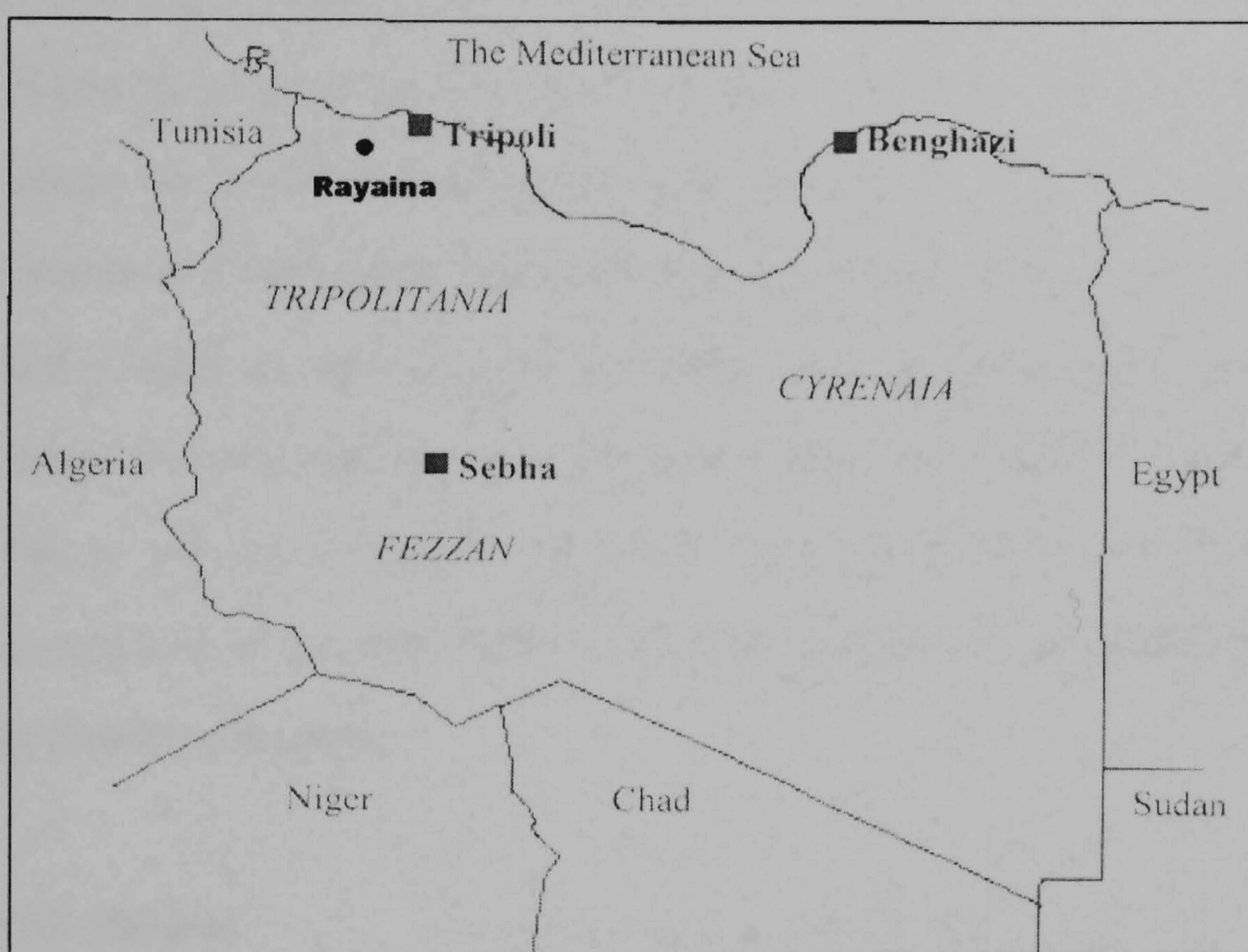
---

[9] https://www.cia.gov/library/publications/the-world-factbook/print/ly.html [Accessed on 05.07.2008].

[10] http://ar.wikipedia.org/wiki/%D8%A7%D9%84%D8%B1%D9%8A%D8%A7%D9%8A%D9%86%D8%A9 [Accessed on 05.07.2008].

The dialect spoken does not exhibit significant differences from others spoken in the region of Tripolitania especially with regard to the phonology of the dialect. For instance the Rayaina dialect has the same number of vowels as other dialects in the region. It should be emphasised that this linguistic information is based on the researcher's experience as a native speaker of the language. The dialect has not been studied before and therefore no relevant references could be found.

Fig.4. 1 A map showing the location of Rayaina where the study dialect is spoken[11]



The dialect exhibits some differences from other dialects as far as its vocabulary is concerned. For example, the words in the left column found in Tripoli dialect (TA) are occasionally realised in Rayaina dialect as those in the column in the middle below.

---

[11] The map was adapted from http://geography.about.com/library/blank/blxlibya.htm [Accessed on 05. 07. 2008].

TA	*RA*	*Gloss.*
[saija:rah]	[karahba]	you want
[awla : d]	[ḍra : ri]	boys
[bui]	[si : di]	my father

However, it should be emphasised that the words found in TA are also used in RA especially among the younger generations thanks to the development of transportation and communication means between the two cities and in the whole country. Tripoli is the capital of Libya and a considerable number of people from Rayaina city commute to it for different purposes such as studying and working. In addition, the media has also contributed to this dialectal effect. TA is the dialect dominant on TV and radio programmes in Libya. Because of this interaction between dialects which has been going on for decades, the new generations of speakers of Rayaina from which the sample of this study is taken are considered to be in regular contacts with the TA dialect. Therefore, their accent can be considered more representative of the wider region which is advantageous for the generalisability of the findings of this study.

4.1.2. Material

The material investigated with the group of participants included sets of words containing the eight Libyan Arabic simple vowels. These are /i :, i, u :, u, e :, o :, æ :, and æ/. It is worth noting here that these vowels have emphatic versions. These emphatic versions, though important, are not dealt with in the study due to time and space limitations. LA diphthongs are also excluded for the same reasons. Moreover, a study contemporaneous to this one is being conducted by a Libyan colleague who is focussing on emphasis in LA (Kriba, forthcoming), in which emphatic vowels will be covered.

Each of the eight vowels was represented in two words on each of five sheets containing the target words. The words were shuffled in a quasi random order so that the two words representing the same vowel on each sheet did not follow each other. Each sheet was recorded twice by each speaker to obtain a total of 20 tokens for each vowel from each speaker.

The data prepared for recording was presented in al-Naskh script on five A4 size sheets, each including the eight target vowels presented in monosyllabic words taking the form of CV(V)C (see Table 4.1) embedded in the carrier sentence /ænæ: gutlæ _____ jæ:æxi/ *(I said to him _____ my brother)* and typed in 28 font size which was thought to be big and clear enough to be read from a reasonable distance without any difficulty (see Appendix 1). The use of a carrier sentence was chosen to prevent unusual word list intonation if these words were produced in isolation and to avoid vowel duration being affected, since vowels tend to be longer when produced in isolated words than when in context (Alghamdi 1990, p.23). Monosyllabic words were favoured over bisyllabic or multisyllabic words for the same reason. Vowels tend to be shorter when occurring in words including a number of syllables than in monosyllabic words.

Efforts were also made to avoid the use of nonsense words, since these have a tendency to be longer than meaningful words with equal numbers of phonetic segments (Umeda 1977). Moreover, nonsense words may not really represent how these vowels are realised in the dialect in question.

In order to help participants avoid pronunciation errors during the recording process, diacritics representing vowels were unusually added where possible to the target words. In MSA, diacritics are used to represent vowels, and therefore those vowels that are unique to the dialect were difficult to represent by diacritics familiar to the participants. In fact, diacritics are only used to represent vowels in MSA for pedagogical purposes or when accurate pronunciation is mandated, as when reading the Holy Koran. Long vowels and consonants, on the other hand, are represented by letters in orthography. The words in table 4.1 are some examples. Table 4.2 shows the words read by the participants and their phonemic transcriptions.

Table 4. 1 The use of diacritics to represent vowels in Arabic orthography

Words without diacritics	*words with diacritics*	*gloss.*
بيك	بِيك	by you
لون	لُون	colour
ساد	سَاد	enough
مد	مِد	hand (v.)
بن	بُن	coffee
مس	مَس	touch

Table 4. 2 Words recorded by participants[12]

Vowels	i:	u:	i	u	æ:	o:	æ	e:
Sheet 1	/bi:k/	/bu:m/	/fin/	/buk/	/læ:m/	/do:r/	/læm/	/be:n/
	/di:n/	/du:m/	/bin/	/bul/	/sæ:d/	/lo:n/	/sæd/	/be:t/
Sheet 2	/gi:s/	/lu:m/	/fiz/	/ful/	/fæ:t/	/lo:z/	/fæd/	/se:f/
	/li:n/	/mu:s/	/hiz/	/bud/	/næ:b/	/fo:z/	/næg/	/se:r/
Sheet 3	/si:d/	/mu:t/	/kis/	/bug/	/bæ:n/	/ko:n/	/sæb/	/de:n/
	/si:n/	/ru:f/	/lim/	/bum/	/bæ:t/	/do:m/	/fæk/	/de:r/
Sheet 4	/ti:n/	/ru:m/	/liz/	/bun/	/sæ:m/	/mo:t/	/ræn/	/ʒe:b/
	/ri:m/	/ru:z/	/mid/	/dub/	/dæ:m/	/ro:z/	/hæm/	/ze:d/
Sheet 5	/ki:s/	/tu:t/	/mis/	/hum/	/fæ:d/	/ħo:ʃ/	/mæs/	/ze:n/
	/li:g/	/bu:k/	/tim/	/kub/	/gæ:s/	/go:s/	/wæn/	/ke:l/

12 A glossary illustration of the meanings of these words can be found in appendix 5.

4.1.2. Sample

The sample included 20 participants, all of whom were from the city of Rayaina. The speakers had lived in the same town where the data was collected since birth. However, these participants had regular contact with other dialects in the region especially TA, the dialect spoken in the capital city of Tripoli, for the reasons mentioned in Section 4.1.1 above. The greater number of participants in comparison to previous studies (e.g., Hussain 1985, Mitleb 1984a, Flege 1979, Al-Ani 1970) was chosen in order to allow for the better generalisation of results.

The choice of participants residing in a place where the language tested is spoken was made in order to ensure that the results are not affected by a foreign language environment. For instance, in their study of voice onset time (VOT) perception in various languages, Keating, Mikos and Ganog (1981) found that Polish listeners who lived in the United States responded differently from those who lived in Poland despite the fact that the participants in the study had very little knowledge of other languages. Similarly, Flege and Eefting (1987) found that monolingual Spanish speakers produce a longer voicing lead for their voiced stops than Spanish speakers who spoke English. In this study, some experience in a foreign language was inevitable because English is part of the taught programme in the Libyan educational system (see Table 4.3 below). All participants were male because it would have been difficult for the researcher to obtain the data from women for cultural reasons.

4.1.3. Recording procedure

The material produced by speakers was recorded using an AudioTechnica ATR25 microphone connected to a Toshiba laptop via an external sound USB Sound Blaster Audigy 2 NX card with a sampling rate of 22050 Hz, 16 bits, mono. The data were saved as wav files on the laptop and copies were also stored on USB flash memories.

Table 4. 3 Participants' age and English experience (production task)

Participants	Age	Foreign language experience in years		
		Preparatory	Secondary	University/college
1	24	3	3	1
2	19	3	3	0
3	21	3	3	1
4	21	3	3	1
5	21	3	4	1
6	44	3	3	0
7	30	0	3	0
8	30	0	3	3
9	22	3	4	2
10	43	3	0	0
11	38	3	3	4
12	32	0	3	1
13	24	3	3	2+ 1 diploma
14	24	3	3	3
15	24	3	3	1
16	25	3	3	1
17	42	2	0	0
18	21	3	3	0
19	19	3	3	4
20	19	3	3	0

The twenty participants in the production task were individually presented with the five sheets containing the target words (as described in section 4.1.1). Each sheet began and ended with words that did not constitute part of the actual data but that were only included to avoid changes in intonation, duration and loudness that often accompany utterances before and after a pause. Because of the length of the list (80 words), the sheets including these words were recorded with short breaks between them in order to avoid tiring the participants and therefore affecting their reading or voice quality. Another reason for the breaks was to allow the participants to reread a particular sheet if any mistakes were made, such as long hesitation, omission, or the wrong pronunciation of words. Participants were given the opportunity to have a look at the words in each sheet before recording and to ask questions about some words to make sure that they were the words intended. Furthermore, each participant was asked to speak normally in terms of loudness and speed. The recordings took place in a quiet room in the city of Rayaina where the participants resided.

Since the target sounds were dialectal, it was emphasised to the participants that they should produce the utterances in their dialect in spite of the fact that the written letters represent the standard variety and some of these words do exist in MSA. When an MSA pronunciation occurred, the recording of that sheet was abandoned and the participant was reminded of the fact that these words are dialectal and he was asked to reread the whole sheet again. It should be noted that the dialect is mainly spoken and does not possess a unique writing system. However, the linguistic aim of the recording was never explicitly stated to the subjects in order for their production to be natural. Moreover, no time limit was set, and practice was allowed so that the subjects could familiarise themselves with the words and with the recording situation. The total number of tokens per participant was 160. This amounted to a total of 3200 tokens for the whole group.

Some difficulties occurred during the recording process. First, some participants were not able to read properly due to their levels of literacy, and they were excluded and replaced by others to achieve the target sample size. Second, although participants were allowed to have a look at the words written before recording and to ask questions, some of them encountered difficulty in reading the words as they are pronounced in the dialect. This was expected due to the nature of the written form and some common vocabulary between LA and MSA as mentioned above. In order to overcome this problem, the researcher repeatedly emphasised to the participants that the words are dialectal and should be pronounced as they are said in the dialect and not in MSA.

In spite of the above mentioned difficulties, eliciting the material produced in the target dialect was fairly successful. This is due to the methods used in eliciting the material from the native speakers of the dialect. Efforts were made by the researcher to solve any problem that might have a negative impact on the results and to make such elicitation methods effective. For instance, it was ensured that highly technical equipment is used for the recording process to guarantee the clarity of sound and the accuracy of measurements. Also, a robust sample was used to allow for more reliability and generalisation of results. Finally, a sufficient number of words that contain all the vowels of the dialect in different contexts was necessary for the same reasons.

To sum up, despite the procedural difficulties mentioned above, the methods adopted were effective in that a suitable sample of speech was generated and which constitutes an appropriate basis for investigating the research questions that this study aims to answer.

4.1.4. Acoustic measurements

The measurements taken included F1 and F2 and the duration of vowels. As Ladefoged points out "vowels can always be accurately described in terms of the frequencies of the first three formants. It is often sufficient to plot the frequencies of the first two formants on a formant chart" (Ladefoged 2001, p.46). Therefore the focus in this study on F1 and F2 is due to their more important role in examining vowel categories and in mapping acoustic realisation to articulatory expectations. However, because F3 mainly plays a significant role in /r/ coloured vowels

(Ladefoged 2001, p.46) and since this type of vowels does not exist in Arabic, the measurements of formant frequencies were limited to the first two formants. Moreover, most of the studies related to other Arabic dialects have concentrated on the first two formants, and for comparative purposes the focus in this study is limited to these two formants.

Before measuring the formants and the durations of each vowel, a phonetic transcription of that vowel was taken (Appendix 5). Then, tokens were measured by hand using the speech analysis software Praat (Boersma & Weenink 2006), which displays speech samples using a digital spectrogram and time domain waveform. Default settings of the program were used in the analysis except when it was necessary to change these settings (see Table 4.4 below). For instance, changes in the amplitude and shape of the time domain waveform were used as a secondary reference for defining vowel boundaries. Formant tracks, which were overlaid on the spectrogram, were also used. However, because the automatic tracker in Praat is not always accurate, visual inspections of spectrograms in addition to the tracker were always made.

When taking the formant measurements and vowel duration several factors were taken into consideration. Measurements of these formants were made at an

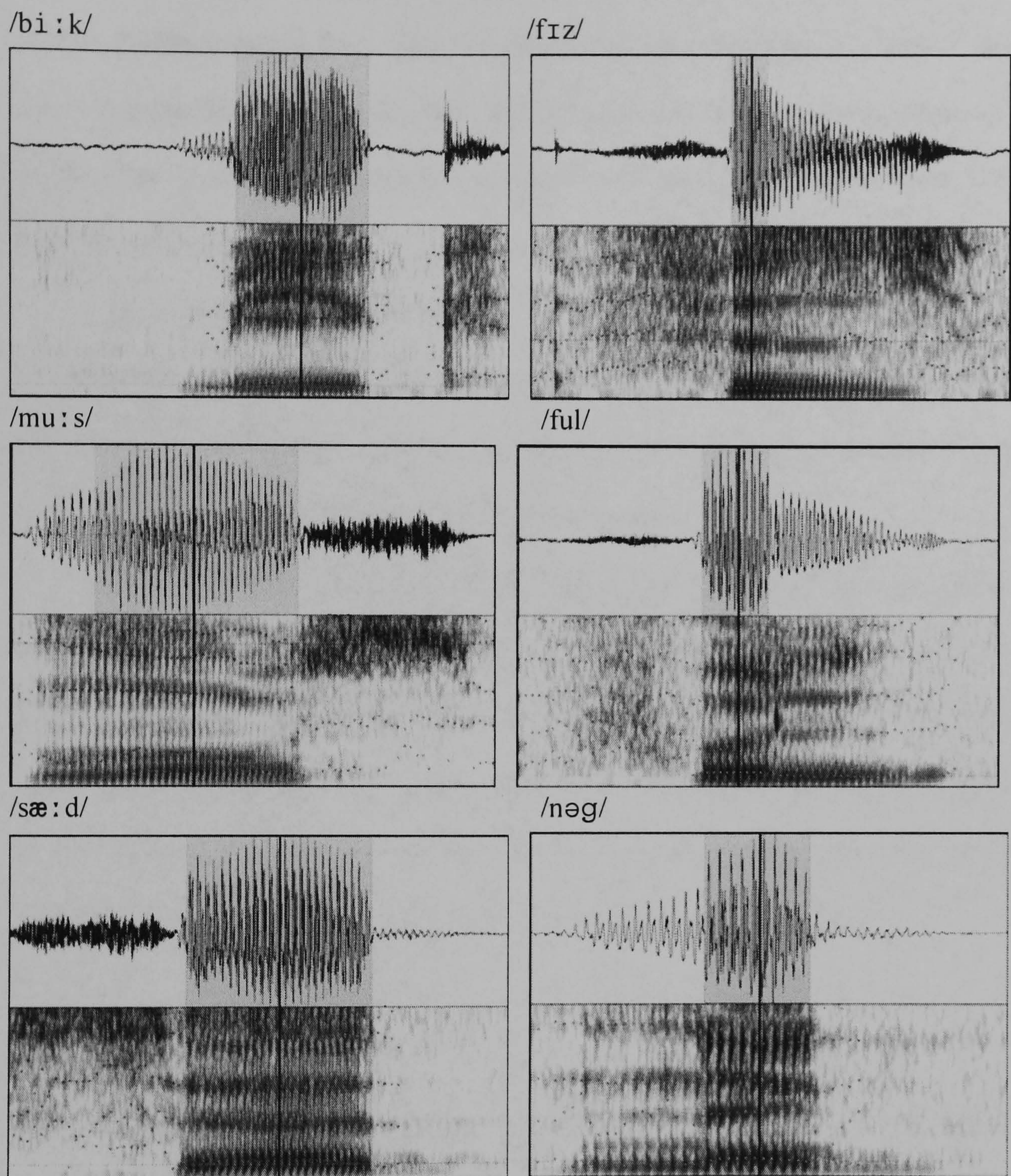
Table 4. 4 The default settings of Praat used in measuring the formants and duration

Spectrogram settings			*Formant settings*				
View range (Hz)	Window length (s)	Dynamic range (dB)	Maximum formant (Hz)	Number of formants	Window length (s)	Dynamic range (dB)	Dot size (mm)
0.0-5000.0	0.005	50.0	5500.0	5.0	0.025	30.0	1.0

interval in the middle of the vowel (following Rosner & Pickering 1994, p.11; Frieda et al 2000, p.133) far enough from the transitions of the consonants at the beginning and at the end of the word (see Figure 4.2 below). This interval is supposed to be the section of the vowel least influenced by neighbouring consonants (Lindblom 1963; Lehiste & Peterson 1961). Ideally, measurements should be made at a point where all formants are steady (Ladefoged 2003, p.105). However, this was not always possible and there were times where the most suitable point for measuring one of the formants was not optimal for measuring the other formant. In such cases some compromise was obligatory and measurements were made at a point that was thought to be most suitable for both formants bearing in mind that the measurement procedure should be relatively consistent across all tokens. All formant measurements were rounded to the nearest full number.

Vowel duration was measured from the onset of energy in F1 to the offset of energy in F1 and F2 which, according to Flege and Port (1981, p.128), mark vowel boundaries. The VOT of the voiceless stop that may precede the vowel was not considered as part of the vowel duration (Petrson & Lehiste 1960, Fant 1958, 1960) (see Figure 4.2 below). Duration was given by Praat in seconds and then transformed and presented in figures, tables and diagrams in milliseconds.

Fig.4. 2 Spectrograms illustrating place of formant and duration measurements of some words as produced by one speaker (The selected areas represent the duration of the vowel while the line in the middle of the selected area refers to where the formants were measured)



4.1.5. Reliability of the measurements

Reliability is mainly concerned with the consistency of the measures; that is, whether the measurements taken are consistent and repeatable (Bryman 2001, p.29). With regard to the production results and in order to check the reliability of the formant measurements, suggestions made by Ladefoged (2003, p.127) were followed.

First, when all measurements were finished, they were repeated. Because of the large number of tokens, it took approximately one month to finish all the measurements and another month to repeat them. The two repetitions were then compared and if there was a substantial difference between the first and the second measurements of a certain formant, a third measurement was made in order to find out which one of the two was more accurate.

Since all tokens were recorded twice, the second step to ensure reliability was to compare the formants of the two repeated tokens. If, there was a difference of more than 50 Hz between the two repetitions of the same formant, they were checked again to ensure that there was no fault with the measurements.

The final step in ensuring the reliability of the measurements was to plot all tokens produced by a certain speaker for a certain vowel on the formant plane. This process would highlight any outliers in the production of that vowel by that speaker. When an outlier was found, the formant measurement of that outlier was taken again. If no mistakes were found, the data were accepted as showing intra-speaker variability. The same procedure was followed with duration measurements.

4.1.5.1. The issue of nasals

As can be seen from Table 4.1, in some of the words under investigation the vowel is followed by a nasal. Nasals are characterised by an F1 that has less energy and lower frequency than that of the preceding (or following) vowel. They are also marked by little energy in the region typically occupied by F2 (Ladefoged 2001, p. 54). It is generally thought that nasals lower the formants of the preceding vowel, making their inclusion in the current dataset undesirable. For example, Whalen & Beddor (1989) observed nasal effects on vowels in the form of a lower F1 of low vowels and a lower L2 of front vowels. The formant shift of a vowel in the context of a nasal was also confirmed by Arai (2004) who found that F1 has a tendency to shift in a more central direction when nasalized in English.

However, the decision to keep these words was necessary due to the need for monosyllabic real words as it was not possible to find alternatives that do not contain nasals. Moreover, since LA does not contain nasal vowels, it was hoped that the effect of lowering the velum to produce nasals would be minimised. Finally, target words with final nasals were included for each of the vowels under investigation, so that any nasal effect would influence all vowels and all speakers under investigation.

In order to further investigate the effect of nasals on the preceding vowel, the high front long vowel /iː/ was considered as an example for further analysis because it is followed by a number of nasals that is equal to the number of non-nasals (i.e., 5 words each).

Table 4. 5 Average formants for /iː/ followed by nasals and non-nasals for each of the 20 speakers

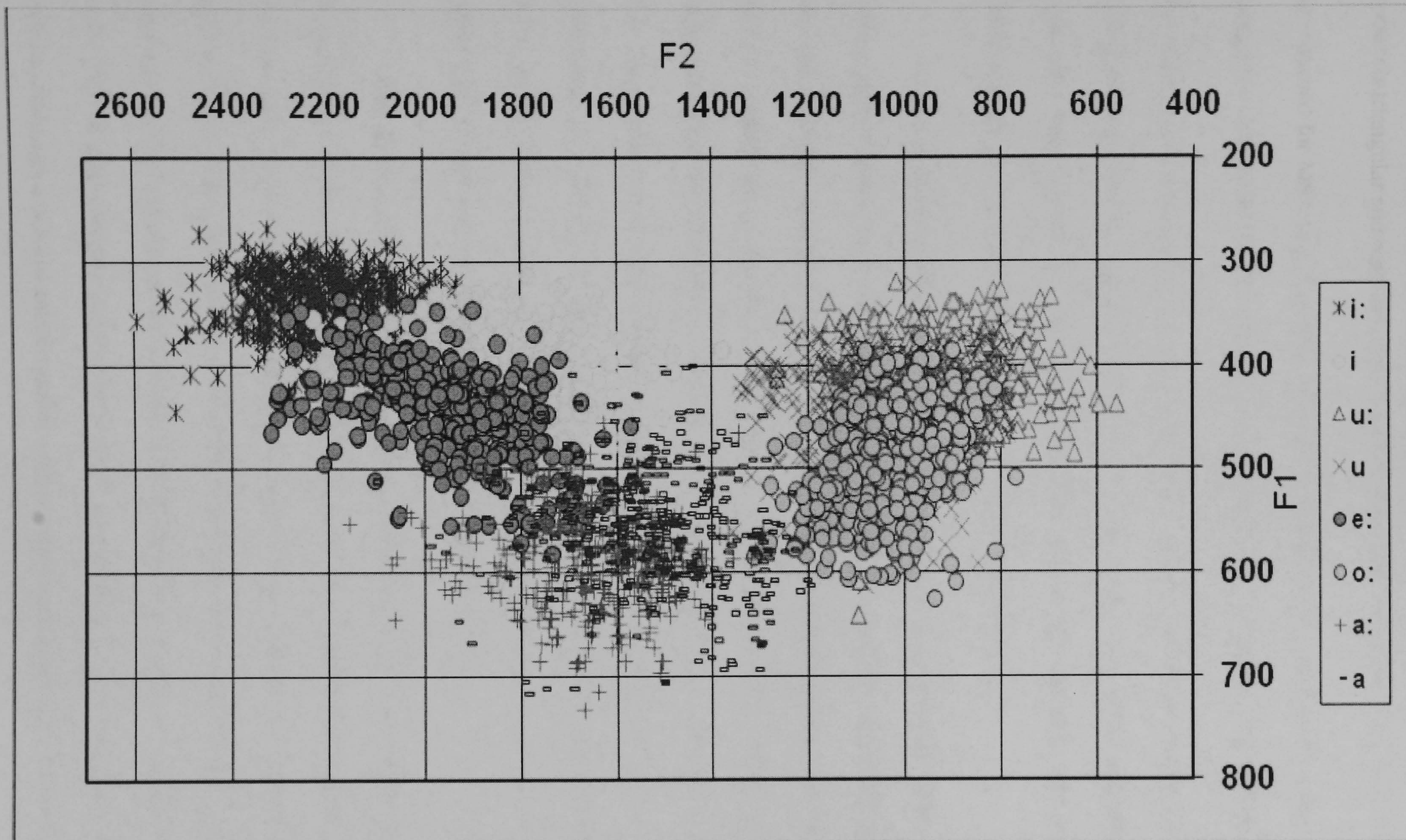
Speaker	F1		F2	
	Nasals	Non-nasals	Nasals	Non-nasals
1	372	358	2296	2282
2	344	325	2343	2340
3	376	327	2005	1973
4	425	380	2138	2045
5	358	317	2256	2234
6	322	328	2376	2296
7	343	316	2241	2191
8	304	299	2196	2208
9	360	323	2419	2321
10	324	311	2050	2122
11	353	333	2195	2151
12	326	319	2056	2096
13	319	310	2237	2194
14	341	369	2279	2259
15	386	361	2241	2180
16	349	340	2163	2158
17	323	322	2302	2276
18	351	319	2398	2319
19	331	334	2117	2121
20	378	386	2243	2208
All speakers	**349**	**334**	**2228**	**2198**

Table 4.5 shows the average mean formants for the vowel /iː/followed by both nasals and non-nasals for the 20 speakers. The Table unexpectedly shows that average formants for most of the speakers for the vowel /iː/ followed by nasals are higher than for those followed by non-nasals. Moreover, when a Paired Sample Test was conducted, this difference was found to be significant for both F1 and F2 ($t = 3.498$, $df = 19$ and p value is 0.002 for F1; $t = 2.996$, $df = 19$ and the p value is 0.007 for F2; shaded cells represent the higher mean values). Only four speakers showed higher formant values for F1 and F2 for nasals than for non-nasals. This might indicate that nasals have little effect on preceding vowels for LA vowels in this study.

4.2.1. Formant analysis results

In this section the overall results of the production of all 20 speakers are presented before looking at individual differences in vowel realisation. First, Figure 4.2 shows the vowel plane representing all tokens by all speakers for the eight vowels that make the LA vocalic sound system. In Figure 4.2, F2 values represented by the horizontal axis are plotted against F1 values which are represented by the vertical axis and both of which originate in the right hand upper corner of the plane in order for the vowel plot to reflect the two dimensional properties of the vowel quadrilateral. This configuration was first suggested by Joos (1948) in order for vowel acoustic analysis to parallel vowel articulation. F1 correlates with tongue height and F2 reflects tongue advancement. According to Rosner & Pickering (1994, p.11) a figure of this type provides a suitable method, though not perfect, of representing a vowel system of a given language.

Fig.4. 3 LA vowels plotted on an F1 vs. F2 plane (Hertz)



As seen in Figure 4.3, the overall acoustic pattern of the Libyan Arabic vowel system is triangular and represented by the three main vowel qualities /iː/, /uː/, /aː/ that are reported for MSA (e.g. Gairdner 1925, p.33; Mitchell 1993, p.138). The other vowel categories seem to follow this triangular shape with the /eː/ and /oː/ slots falling along the sides of the triangle and the short counterparts of the three main vocalic elements falling on the inside. The degree to which short vowels appear to be centralised compared with their long counterparts varies between short vowels. For instance, this is more obvious in the case of /i/-/iː/ than between /u/ and /uː/ or /a/ and /aː/.

It is also apparent from the formant plane that there is considerable intra-vowel scattering and inter-vowel overlap to the degree that some vowels are almost included in the space of other vowels. Although formant frequencies have long been seen as main factors in determining the phonetic quality of vowels, the absolute values of these formants are greatly affected by differences within as well as between speakers. Some of this inter-speaker and intra-speaker variation can be attributed to factors such as differences in vocal tract length and contextual differences (e.g., Rosner & Pickering 1994, p.15). These differences result in considerable variation in the production of speakers when formants are plotted on a plane (Disner 1986, p.69).

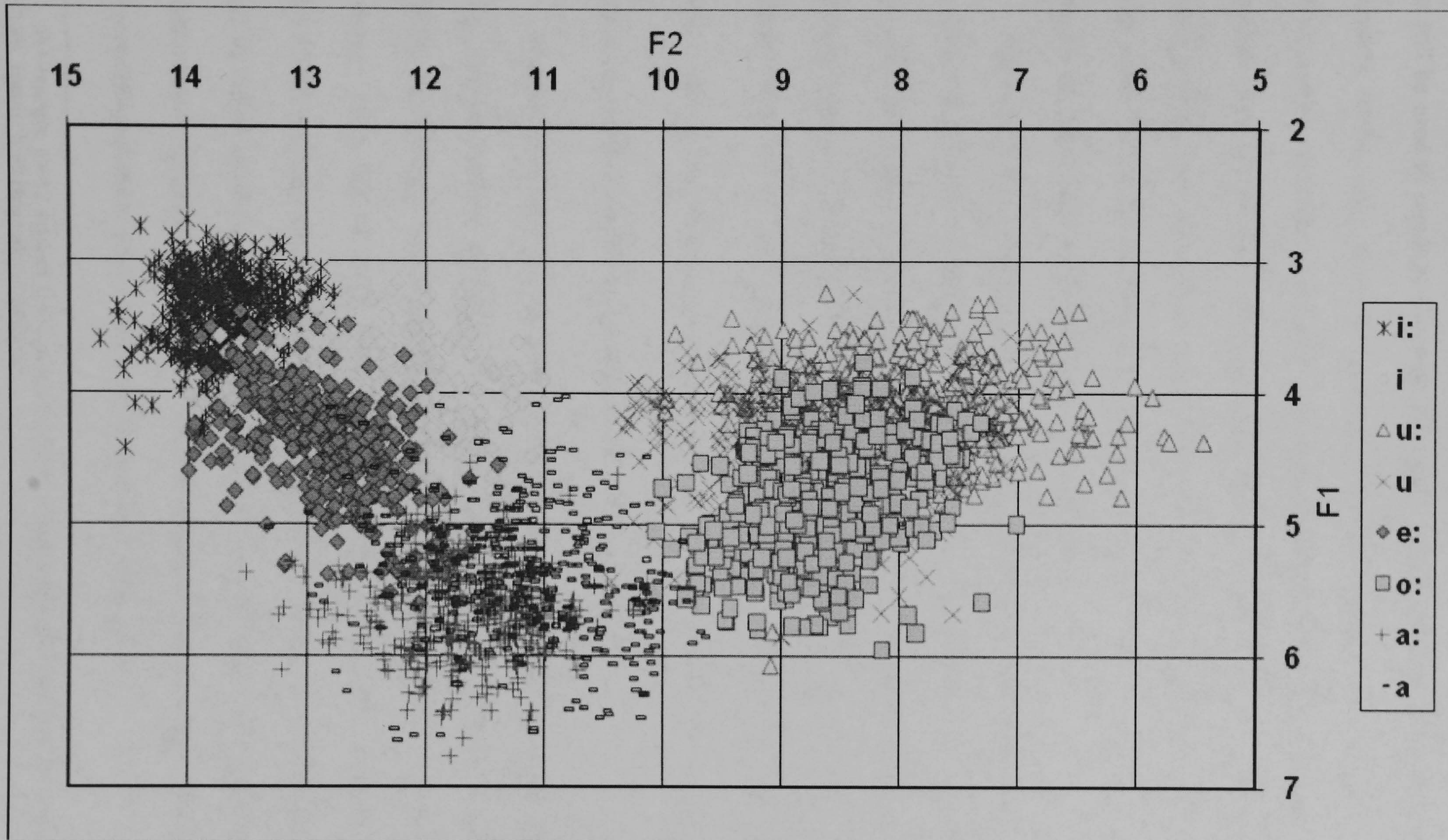
This vowel plot representing the production of all speakers showed that speech is considerably variable both within and between speakers. Phonologically identical vowels were found to show acoustic variation across speakers. This variation can be attributed to several factors such as the context in which these vowels occur and to differences in vocal tract length. Yet, listeners are still capable of recognising these vowels as belonging to the same phonological category despite the variation they exhibit in the vowel space. This is achieved through a process called speaker normalisation (see Section 2.5.1.3 in Chapter

2). Research has long been concerned with how this process happens; that is, how listeners are able to identify vowels while their relevant acoustic cues are not exactly identical. One explanation for this is that when a listener hears an utterance, they construct a complete vowel pattern to be used as a reference system against which they locate new vowels when they hear them (Joos 1948, p.16). In doing this and in addition to the information internal to the vowel, the listener also exploits other sources of information such as those related to the speech context and visual information about the speaker in order to specify their vocal tract size.

According to Johnson et al. (1993), vocal tract size differences are not only found between- but also within-genders. Converting formant values in Hertz to Bark values is commonly used in order to represent vowels using an auditory scale rather than linear measures of frequency in Hertz. The Bark scale is used to represent sensitivity to frequency differences in critical bands. This is thought to be similar to listeners' behaviour which demonstrates that they are more sensitive to changes of lower frequency than higher frequency in Hertz (Stephenson 2004, p. 366). Therefore, the Bark scale is used to represent the ability of the listener's ear to make a distinction between tones at different frequencies (Zwicker 1961). For instance the human ear is more sensitive to tonal differences between 1000 Hz and 2000Hz than between 4000Hz and 5000Hz. By using the bark scale the vowel space can be stretched where the ear is most sensitive to tonal differences or contracted where the tonal differences are hard to distinguish (Aylett , p. 2).

In order to minimise the effect of non-linguistic factors, including the different shapes of the vocal tract that participants may have, and to allow for better comparison between different speakers, formant frequencies in this study were transformed to the bark scale using Traunmuller's (1990, p. 99) method ($b = 26.81/(1+(1960/f)) - 0.53$) (f is in Hz, b is

Fig.4. 4 LA vowels plotted on F1 vs. F2 formants chart (Bark)



in Bark)[13] and represented again in Figure 4.4. For the above mentioned reasons the bark scale will be used to represent speakers' production in scatter plots. The bark scale is particularly useful when differences among individual speakers are considered. As mentioned above, the scale considerably eliminates differences in formants resulting from differences in vocal tract size between speakers. However, because the use of Hertz is so common among researchers, Hertz values will also be used. This will be particularly helpful when the dialect is compared with other Arabic dialects studied by other researcher who have mainly used Hertz in their data analysis.

Although it better represents the perceived distances in the phonetic space, the bark scale plane still shows variations in the production of all vowels by all speakers. Such variations might have resulted from various factors including context and idiosyncratic patterns (Frieda 2000, p.140). The effect of these factors is obvious in the production of all vowels. For example, while vowel /eː/ shares the space with the short vowel /i/ and the long vowel /iː/, vowel /oː/ is almost included in the area occupied by the vowel /u/ which, in turn overlaps with its long counterpart /uː/.

The formant patterns in this study confirm the fact that vowels in natural speech overlap and that formants are not the only distinguishing factors between them. Speech variability has frequently been reported in the literature of language research (e.g. Petrson and Barney 1952, Perkell 1990, Al-Tamimi et al 2002). For example, Peterson and Barney (1952) found considerable inter-speaker variability in the production of American vowels by native speakers. The vowel space occupied by tokens of a particular vowel type produced in different contexts has often been found to overlap significantly with areas containing tokens of other vowel type (Strange et al 1983, p.695).

[13] For example, the F2 value of 1669 can be transferred to Bark as follows: 26.81/(1+(1960/1669)) - 0.53=11.80. That is, 1669 Hz equals 11.80 Bark.

Regarding the effect of vowel length on the distribution of vowels in the vowel space, the F1/F2 plots of short vowels /i, u, a/ and their long counterparts /iː, uː, aː/ seem to suggest that short vowels are more centralised than their long counterparts, as can be seen from the representation of the two vocalic groups on the formant planes shown in Figure 4.3. Assuming there is a correlation between acoustic and articulatory plots, the effect this has on each vowel is the following: the short vocalic category represented by /i/ seems to be more centralised and lower than its long counterpart /iː/. In contrast, the short vowel /u/ is somewhat lower and more to the front than its long counterpart /uː/. The short vowel /a/, on the other hand, is higher and more retracted than its long counterpart /aː/. As mentioned in section 3.1.1 in Chapter Three, there are two main points of view regarding the qualitative contrast between Arabic short vowels and their long counterparts. The traditional view maintains that these vowels are only different in quantity while others claim that these vowels also contrast in their quality as well as quantity. Within the latter view, some scholars have found a significant qualitative difference between all short vowels and their long counterparts while others have found that this difference is apparent in some but not all of these vocalic pairs. For example, Al-Ani (1970) found comparable formant values for front and back vowels. In fact, he found no difference between the means for F2 /iː/ and /i/ (2200 Hz) and a small but insignificant difference in F1 means (290 Hz vs 285 Hz). The same small difference was found between /uː/ and /u/ for both F1 and F2 (285 Hz vs 290 Hz and 775 Hz vs 800 Hz). However, this difference was greater in the case of /aː/ and /a/ (675 Hz vs 600 Hz for F1 and 1200 Hz vs 1500 Hz for F2) (see Table 3.1 in Chapter Three). On the other hand, Newman and Verhoeven (2002) found the opposite situation in CA. That is, the difference between /i:/ and /i/ and between /u:/ and /u/ was greater than that between /a:/ and /a/. The latter two differed by only 4 Hz for F1 and 5 Hz for F2.

Another example that showed differences between all short vowels and their counterparts is Jordanian Arabic. Abou-Hiadar (1994) found a big difference between /i/ and /i:/ (245 Hz for F1 and 575 Hz for F2) and between /u/ and /u:/ (320 Hz for F1 and 445 Hz for F2). The situation in Jordanian Arabic for the low vowels /a/ and /a:/ was different. It was found that the long vowel was higher by 10 Hz and more retracted by 99 Hz than its

Table 4. 6 Differences in Hz between long vowels and their short counterparts in some Arabic dialects.[14]

Vowel	i:-i		u:-u		a:-a	
Dialect	F1	F2	F1	F2	F1	F2
Libyan (this study)	62	358	27	119	33	100
Tunisian (Abou Haidar 1994)	195	585	180	305	-40	190
Tunisian (Belkaid 1984)	70	365	30	205	25	80
Iraqi(Al-Ani 1970)	5	0	5	25	75	-300
Syrian (Abou Haidar 1994)	85	330	110	580	10	-120
Jordanian (Abou Haidar 1994)	245	575	320	445	-10	-99
Saudi (Alghamdi 1998)	110	445	101	344	82	50
Saudi (Abou Haidar 1994)	235	700	165	260	35	-50
Sudanese (Abou Haidar 1994)	95	220	75	140	70	-100
Sudanese(Alghamdi 1998)	59	189	35	324	110	-72
Egyptian (Alghamdi 1998)	101	426	51	343	-6	172
Cairene (Newman 2002)	85	365	70	82	-73	65
Qatari (Abou Haidar 1994)	190	590	180	175	1	-260
Lebanese (Abou Haidar 1994)	210	480	145	265	-30	40
Emirati (Abou Haidar 1994)	125	345	125	85	90	-280

14 Negative values in F1 indicate that the long vowel /a:/ has a lower F1 than its short counterpart and the negative values in F2 indicate that the long vowel /a:/ has a lower F2 than its short counterpart.

short counterpart. Other studies also showed this trend for F1 (e.g. Tunisian and Cairene) and for F2 (e.g. Iraqi and Sudanese).

Table 4.6 shows differences in Hz between long vowels and their short counterparts for some Arabic dialects, most of which were reported by Newman & Verhoeven (2000). Average formant values for these dialects are shown in Appendix 4.

The findings of this study seem to support the view that considerable qualitative differences occur among all short vowels and their long counterparts as can be seen from both the auditory analysis (see Appendix 5) and the formant values of these vowels in Table 4.7 below.

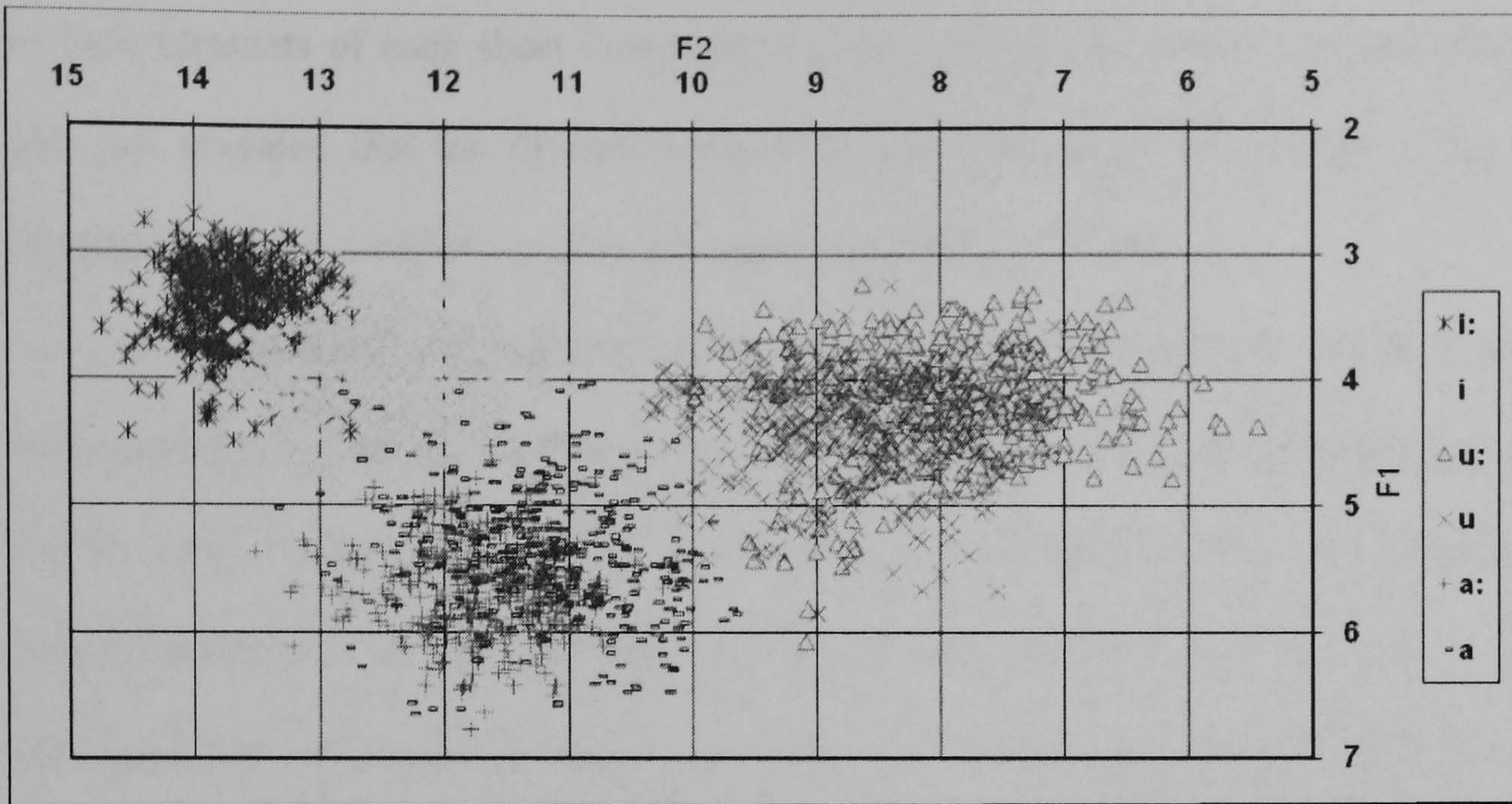
Table 4. 7 Formants mean values of LA long vs short vowels and their difference.

	/iː/	/i/	/uː/	/u/	/aː/	/a/
F1	342	404	416	443	588	555
Difference	**62**		**27**		**33**	
F2	2214	1856	907	1026	1641	1541
Difference	**358**		**119**		**100**	

The formant values in Table 4.4 above also show that the high long vowels /iː and uː/ may not be equally high. This is also true of the high short vowels /i/ and /u/. Figure 4.5 shows the distribution of these vowels in the vowel space as produced by the 20 speakers who participated in the speech production experiment.

When these vowels were examined auditorily (Appendix 5), the quantitative difference between long and short vowels was obvious in all speakers in spite of differences between them in their speech rate. When the focus shifted to qualitative

Fig.4. 5 Distribution of long vowels vs short vowels in the acoustic vowel space (Bark)



differences, it was apparent that the vowel /iː/ and /i/ also differ considerably in their quality in addition to the difference they have in quantity. This difference in quality was also audible in the case of /aː/ and /a/, though to a lesser degree. However, this qualitative difference was not obvious in the case of /uː/-/u/. To minimise contextual effects that may affect this auditory assessment, some minimal pairs found in the data underwent additional investigation. These are shown in Table 4.8 below. These minimal pairs confirmed such difference in quality between short and long vowel.

Table 4. 8 Auditorily compared minimal pairs containing long and short vowel contrasts

/iː/-/i/	/uː/-/u/	/aː/-/a/
/kiːs/-/kis/	/buːk/-/buk/	/saːd/-/sad/ /laːm/-/lam/ /faːd/-/fad/

A Two Sample (Kolmogorov-Smirnov) Test was conducted to compare the means of both formants of each short vowel and its long counterpart as well as their duration. The test revealed that the difference between long and short vowel pairs is not only significant for the duration but also for both F1 and F2 ($p < 0.001$).

Having found a significant qualitative difference between short vowels and their long counterparts, the use of different symbols for these vowels by some researchers (e.g. Cowan 1970, Huthaily 2003) which was mainly based on impressionistic speculations can now be advocated in this study following the experimental analysis of LA vowels. Thus, the high front vowels may be better symbolised by /iː/ and /ɪ/, back vowels as /uː/ and /ʊ/ and, finally, low vowels may have /æː/ and /ə/ symbols. However, it should be emphasised that although the LA /ə/ might not be as central, high and short as its IPA counterpart, it is still the most suitable symbol that might reflect the fact that this short vowel has lower F1 and F2 than its long counterpart /æː/.

Moreover, there might be a suggestion from the production results that these symbols should be used without reference to length by using the length mark (ː) which has been frequently used to differentiate between long vowels and their short counterparts; vowels can still be differentiated by their quality without the need to make reference to their quantity difference in their symbols. However, the perception experiment will help determine which of these acoustic cues plays a major role in identifying these vowels and then a decision to eliminate the use of the length mark (ː) after long vowels might be made.

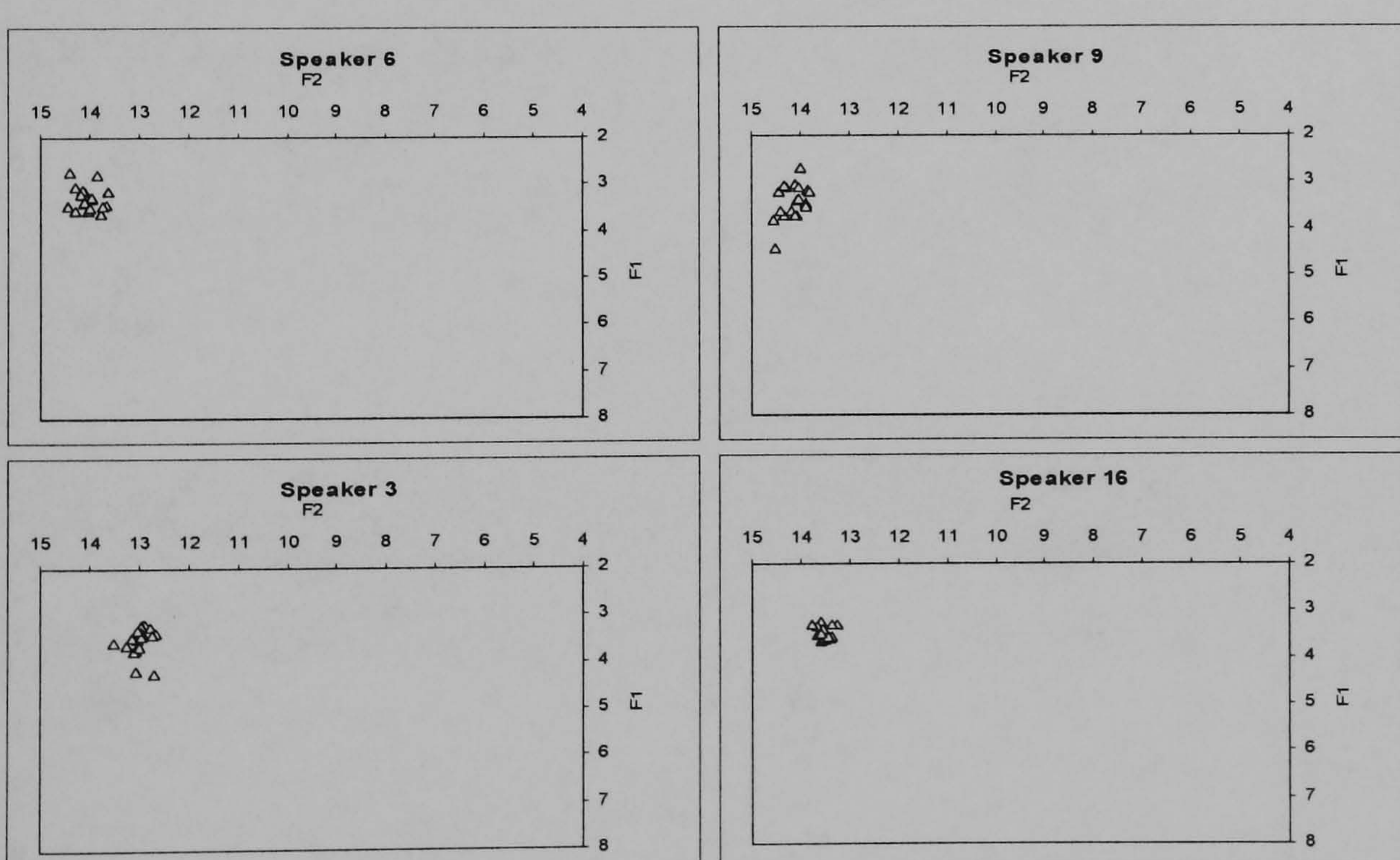
4.2.1.1. Vowel description

Having presented the overall picture regarding the realisation of LA vowels by all speakers, we now move to a more detailed account of how each vowel was realised.

Vowel /iː/

This is a high front long vowel as in /diːn/ *(religion)*, /siːd/ *(master n.)*, /kiːs/ *(bag)*. By looking at the realisation of this vowel by individual speakers (see Appendix 2), it can be noticed that there is considerable variability in their production. Some speakers, for example, speaker 6 and speaker 9, tend to produce it in a more front position in the vowel space than others like, for example, speaker 3 and 16 as appears in Figure 4.6 below.

Fig.4. 6 Acoustic realisations of the vowel /i:/ by some speakers (Bark)

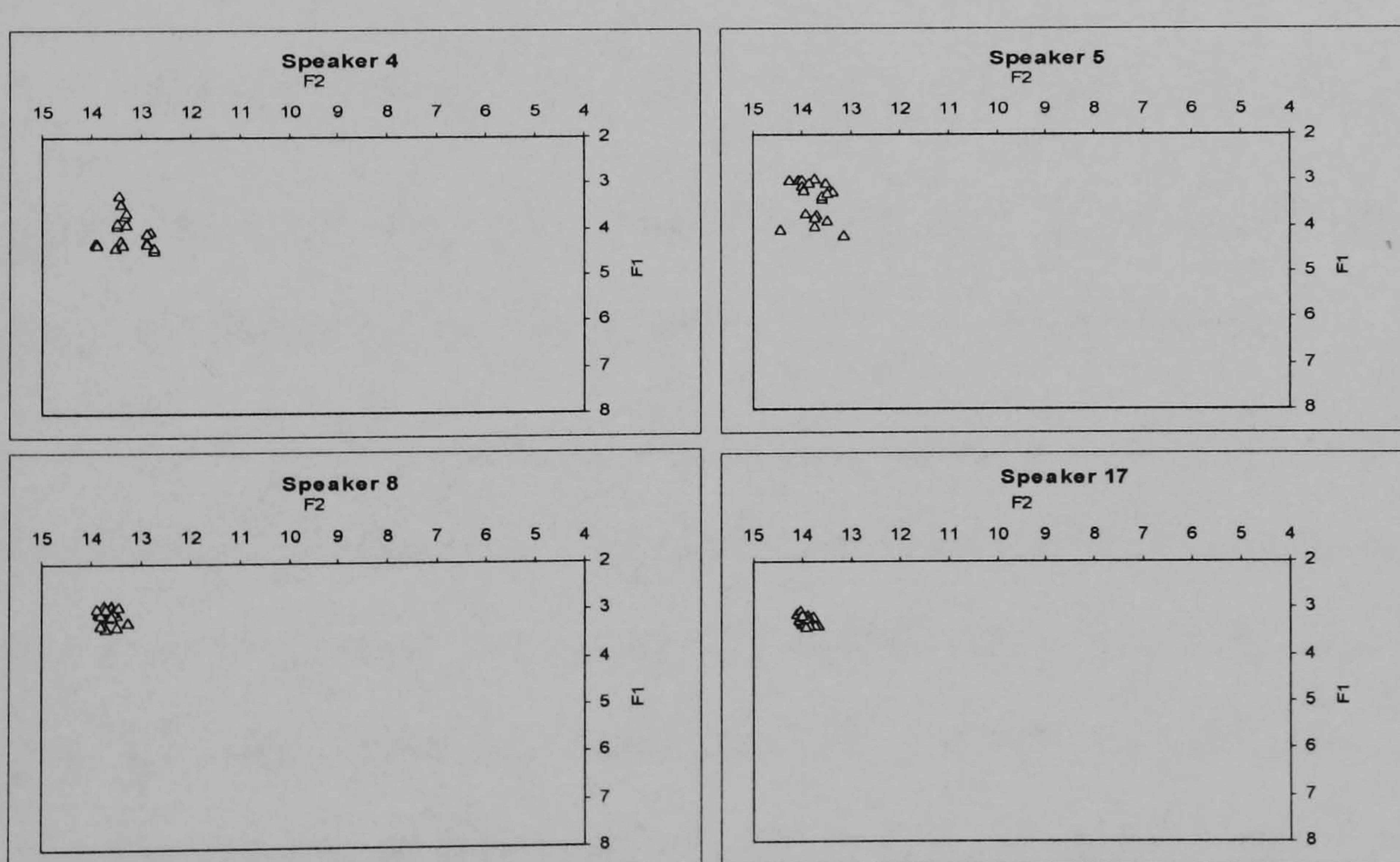


This observation was confirmed when the production of these speakers was averaged. The F2 means for speakers 6 and 9 were 14.05 Bark (2336 Hz) and 14.16 Bark (2374 Hz) respectively while those for speakers 3 and 16 were only 12.97 Bark (1988 Hz) and 13.53 Bark (2160 Hz) respectively.

With regards to vowel height, on the other hand, all speakers tend to exhibit a similar pattern in that their F1 values for this vowel tend to be in a smaller range compared to that of F2 (see Table 4.7).

As for variation within speakers, it can be noticed that vowel realisations are somewhat spread in the acoustic vowel space for some speakers like, for example, speakers 4 and 5 while these realisations tend to cluster together when produced by other speakers like speakers 8 and 17. Figure 4.7 shows variations among these speakers in producing this vowel.

Fig.4. 7 Variation in the acoustic realisation of the vowel /iː/ by some speakers (Bark)



Descriptive statistics of the extent of variation in the production of the vowel /iː/ is presented in Table 4.9, which shows the minimum and maximum values for F1 and F2 in addition to the range between the two values for each formant. The means and the standard deviations are also given.

These descriptive statistics show that the range and the standard deviation in the case of F2 is considerably higher than that of F1 which may indicate that this vowel is less variable along the height dimension than along the back-front dimension.

Table 4. 9 Statistical results for the vowel /iː/

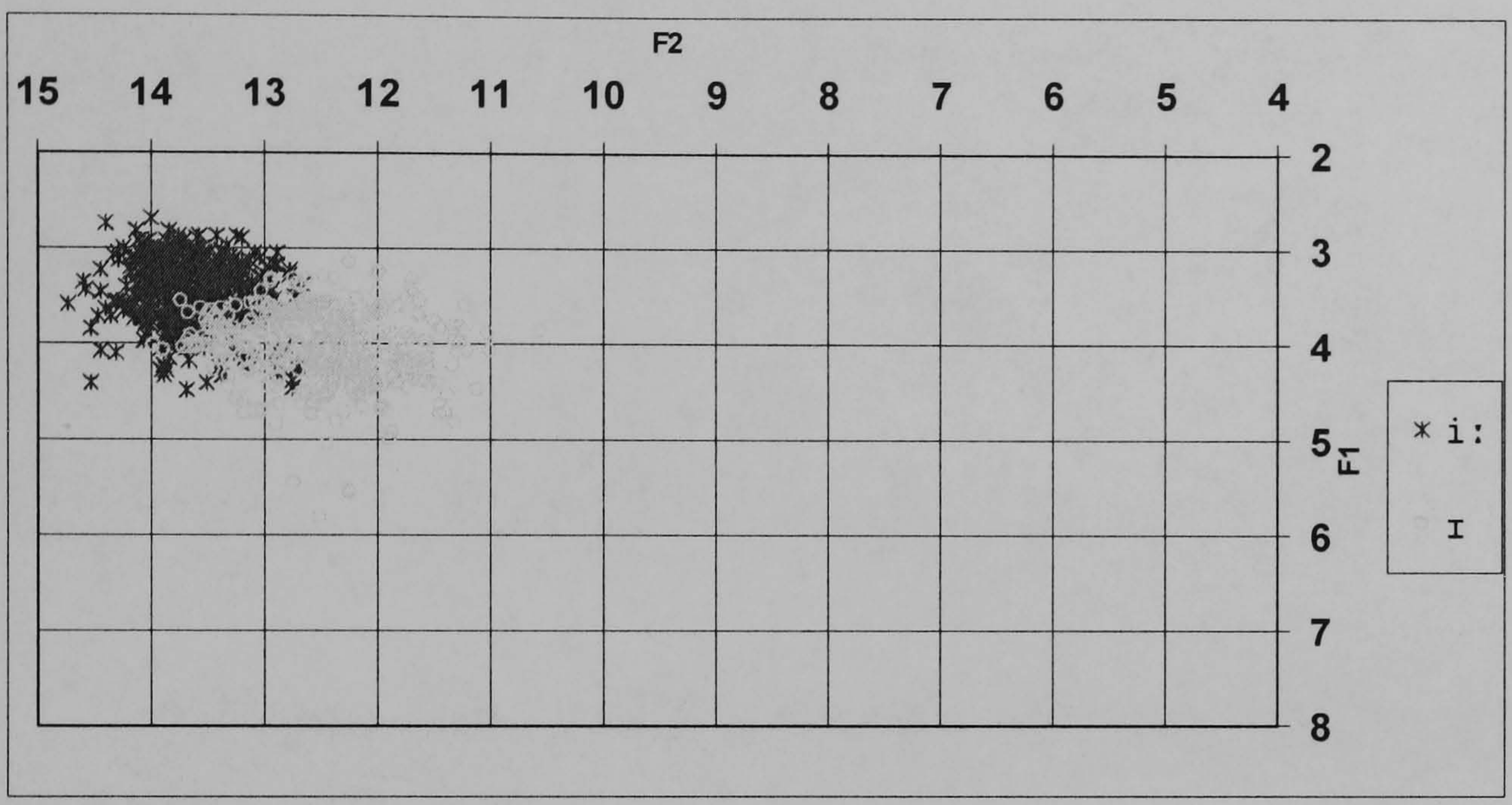
	Maximum	Minimum	Range	Mean	SD
F1	452	268	184	342	35.07
F2	2592	1899	693	2214	125.35

Vowel /ɪ/

This is a high front short vowel as in /bɪn/ *(son of)*, /mɪs/ *(touch)* and /hɪz/ *(shake)*. There is strong evidence from the realisations of this vowel by the 20 speakers that it is retracted and lower than the long vowel /iː/ (see Figure 4.8 below).

A two sample Test (Kolmogorov-Smirnov) shows that the difference between the two vowels is significant at $p < 0.001$ for both formant frequencies as well as duration. This is not in agreement with the general traditional view held by some researchers that these two Arabic vowels only significantly differ in length, especially when MSA is described (e.g., Al-Ani 1970, p.23) (see section 3.1.1.1 in Chapter Three). This suggests that the quality of this vowel in LA may be different from that of MSA or that this vowel may not have been described accurately in the standard variety of Arabic by some scholars. A third possibility would be that this vowel has undergone considerable change in vernacular varieties during the last decades; a possibility that is beyond the scope of this study.

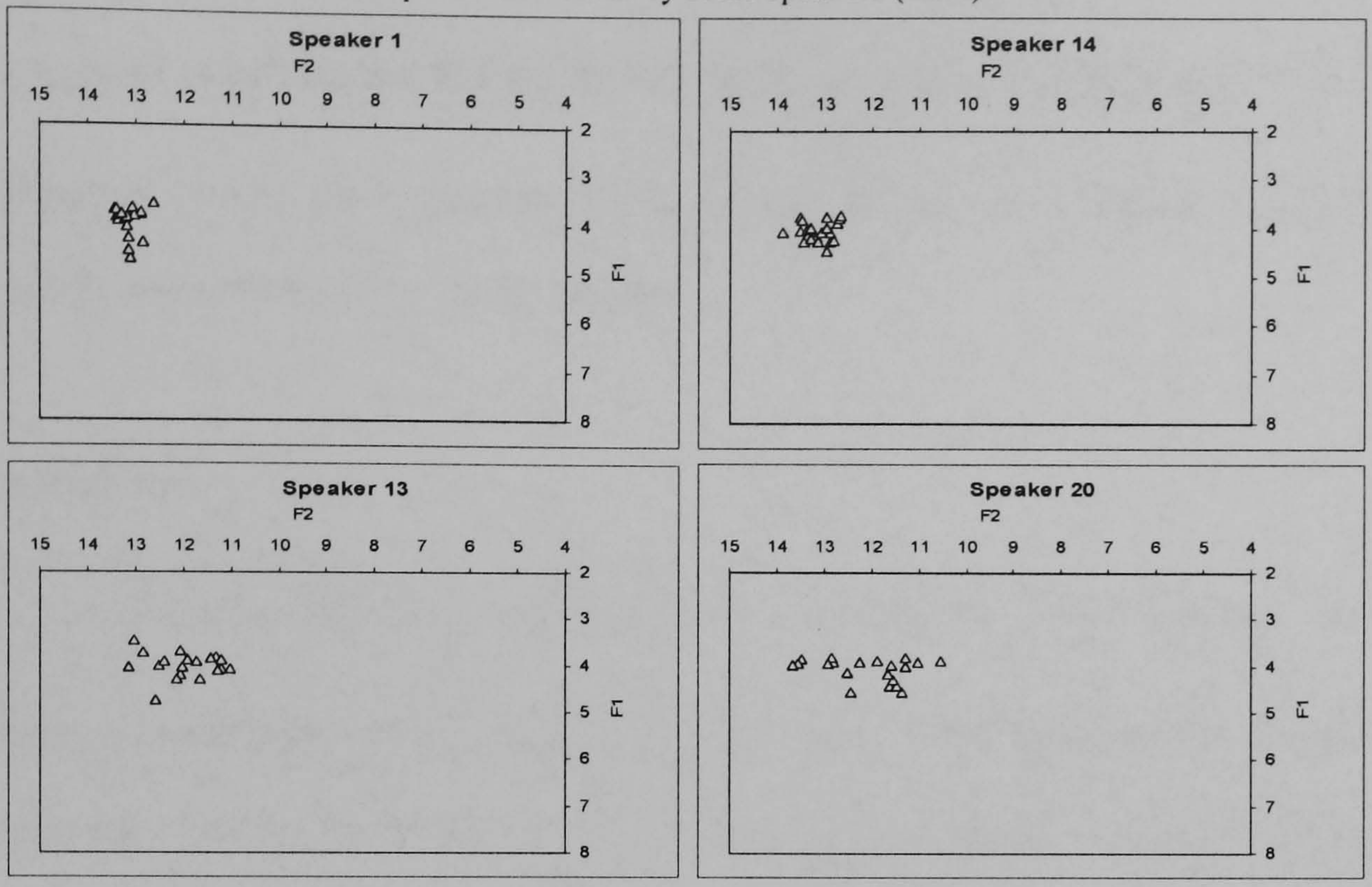
Fig.4. 8 Acoustic realisations of /i:/ and /ɪ/ by all speakers (Bark)



Apart from Al-Ani's study which showed almost no difference in quality between these two vowels, all other studies of several Arabic dialects have shown this qualitative difference between /i:/ and /ɪ/. This difference ranges from 59 Hz for F1 found in Sudanese to 245 Hz found in Jordanian for F1 and from 198 Hz also found in Sudanese to 700 Hz found in Saudi for F2 (Abou-Haidar 1994) (see Table 4.1 above).

The short /ɪ/ category also tends to show more variation in acoustic realisation than that of the long vowel /i:/. Although the tokens of some speakers, like speakers 1 and 14, tend to cluster closely around the mean, others, on the other hand, show a wider spread as in the case of speakers 13 and 20 as Figure 4.9 below shows.

Fig.4. 9 Spread vs. clustered production of /ɪ/ by some speakers (Bark)



This variation in the production of this vowel is further illustrated by Table 4.10 which shows some descriptive statistics on formant values.

Table 4. 10 Descriptive statistics for formant realisations of the vowel /ɪ/ (Hz)

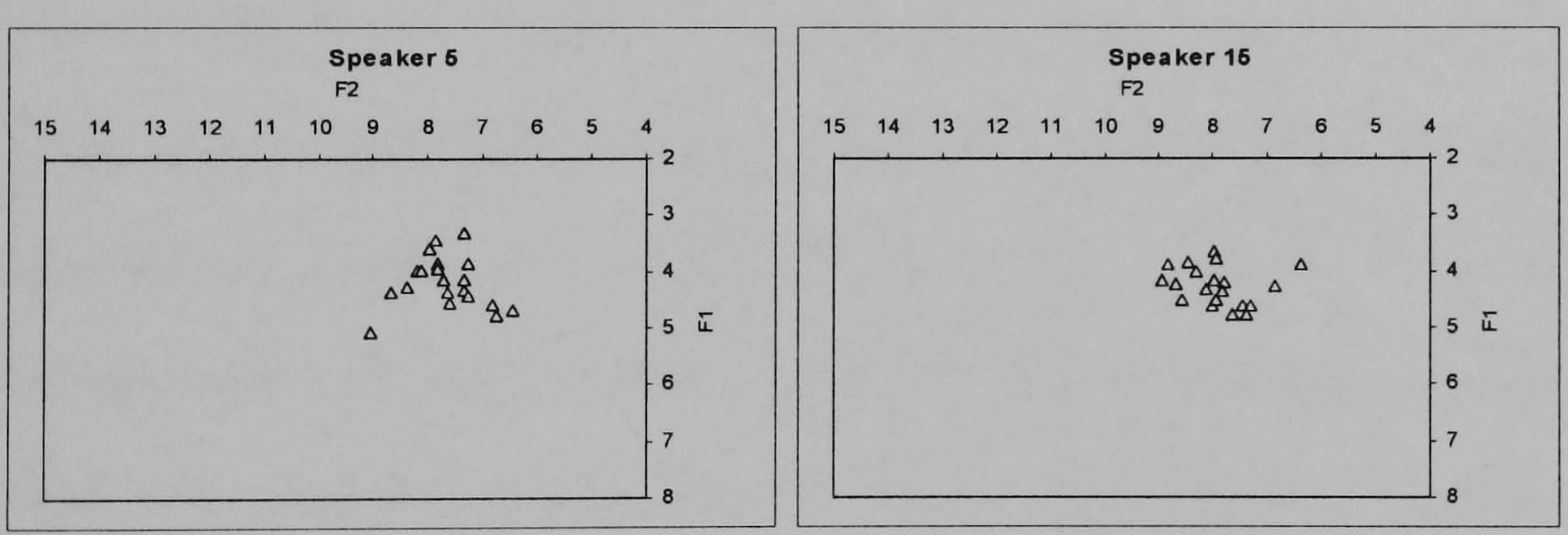
	Maximum	Minimum	Range	Mean	SD
F1	609	313	296	404	41.97
F2	2315	1380	935	1856	164.34

A comparison of these descriptive statistics for the short vowel /ɪ/ with those in Table 4.9 for the long vowel /iː/ confirms the qualitative difference between the two vowels. For example, the F1 mean for /iː/ is 342 Hz compared to 404 Hz for /ɪ/. On the other hand, the F2 mean is 2213 Hz for the long vowel compared to 1855 Hz for the short one. This means that the latter has a tendency to be centralised in the vowel space. As for the range, it is 184 Hz and 693 Hz for F1 and F2 respectively for the long vowel /iː/ compared to 296 Hz and 935 Hz for F1 and F2 respectively for the short vowel /ɪ/. This indicates that the short vowel has more variation in the vowel space than does the long vowel when produced by the 20 speakers.

Vowel /uː/

This is a high back long vowel as in /muːs/*(knife)*, /ruːf/ *(be kind)*, and /buːk/ *(your father)*. This vowel shows wider variation within speakers as well as between speakers than its previously discussed front counterpart /iː/. Compare, for instance, Figure 4.6 for /iː/ with Figure 4.10, which shows /uː/ realisations by some speakers.

Fig.4. 10 Acoustic realisations of the vowel /uː/ by speakers 5 and 15 (Bark)



An examination of the descriptive statistics in Table 4.11 shows that /uː/ exhibits more variation than the long front vowel /iː/ with regards to both F1 and F2. It has a range of 321 and 704 Hz compared to 184 and 693 for F1 and F2 respectively for the high front vowel /iː/.

Table 4. 11 Descriptive statistics for formant realisation of the vowel /u:/ (Hz)

	Maximum	Minimum	Range	Mean	SD
F1	641	320	321	416	49.11
F2	1265	561	704	907	125.24

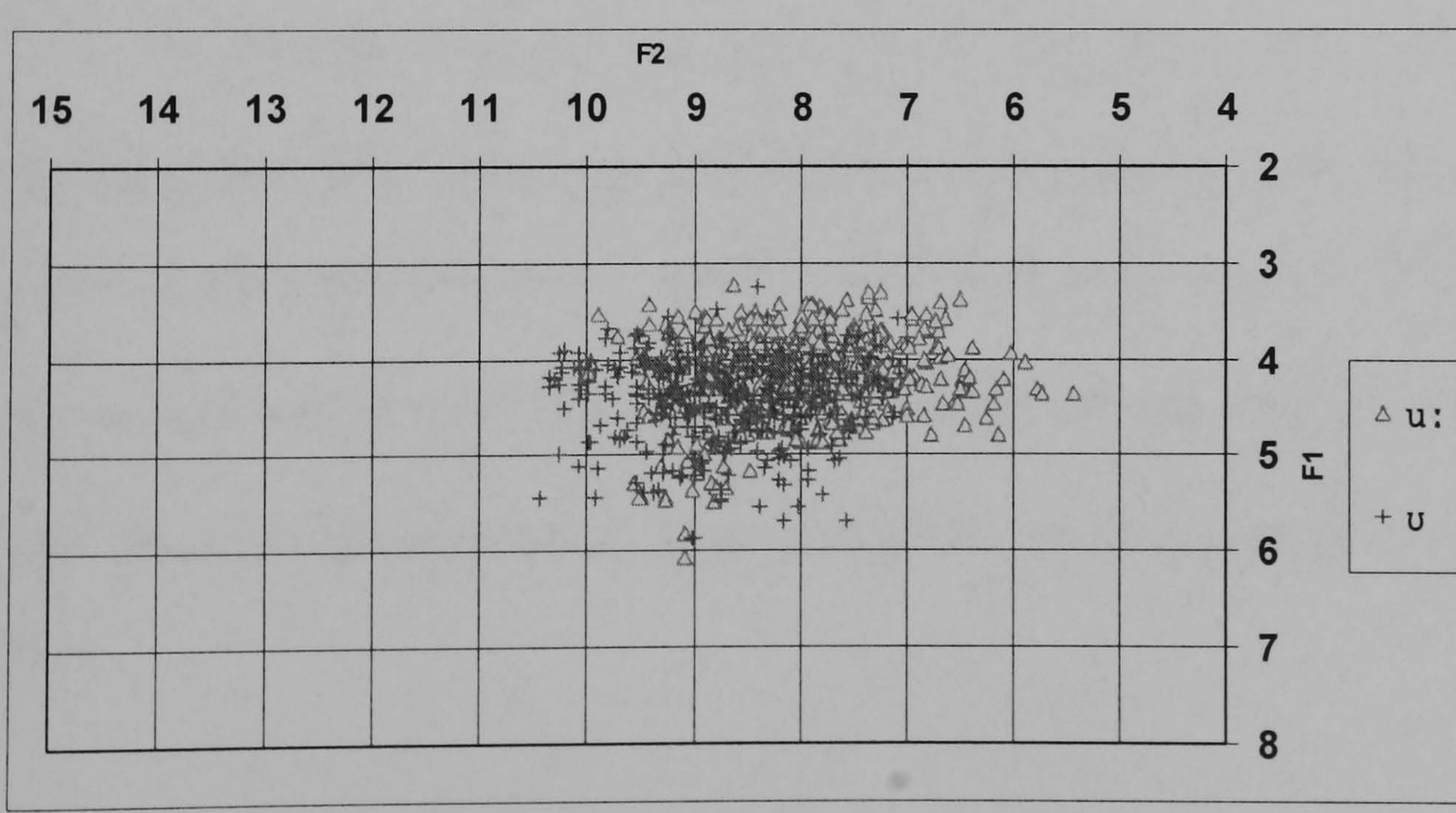
When compared to other Arabic dialects, the mean F1 value for /u:/ is the highest, which suggests that this vowel is lower in the vowel space in LA than in the other Arabic dialects studied so far. The nearest F1 value to the one found in LA is Sudanese which is 380Hz (Abou Haidar 1994). The lowest value, on the other hand, is found in Jordanian which is 260 Hz (Abou Haidar 1994).

The previously made observation that this vowel is lower than its high front counterpart is also evident from the F1 means of the two vowels: 416 Hz for /u:/ compared with 342 Hz for /i:/. The fact that this high back vowel /u:/ is lower than its front counterpart /i:/ was also reported in other Arabic dialects including Tunisian (Abou Haidar 1994; Belkaid 1984), Saudi (Abou Haidar 1994; Alghamdi 1998), Sudanese (Abou Haidar 1994; Alghamdi 1998), Egyptian (Alghamdi 1998), and finally Lebanese and Emirati (Abou Haidar 1994). On the other hand, some other Arabic dialects reported no difference in height between the two high vowels /i:/ and /u:/. These include Iraqi (Al-Ani 1970), Cairene (Newman 2002) and Qatari (Abou Haidar 1994). However, there are two Arabic dialects in which the high front vowel /i:/ is reported to be lower than its back counterpart /u:/. These are Syrian and Jordanian (Abou Haidar 1994). It should be admitted, however, that the difference is only 10 Hz for the former dialect and 40 Hz for the latter. (See Appendix 4 for average formant values for all these dialects.)

Vowel /ʊ/

This is a high back short vowel as in /bʊl/ *(stamp)*, /bʊn/ *(coffee)*, /hʊm/ *(they)*. Compared to the long vowel /u:/, this short vowel does not differ much with regard to its distribution in the vowel space apart from the fact that /ʊ/ tends to be realised with higher F2 and F1, which suggests that this vowel is more front and lower in the vowel space (see Figure 4.11). When a Two Sample (Kolmogorov-Smirnov) Test was conducted to compare the means of short vowels and their long counterparts, the results showed that the difference between /ʊ/ and /u:/ is significant at $p < 0.001$ for both formants and duration. As mentioned above, this tendency for /ʊ/ to be central is shared by all three short vowels in LA.

Fig.4. 11 Acoustic realisations of /u:/ and / ʊ/ by all speakers (Bark)



Some descriptive statistics related to this vowel are shown in Table 4.12 which, when compared to those in 4.11, show the difference between the acoustic realisation of /ʊ/ and /uː/ (444 Hz vs 416 Hz and 1018 Hz vs 907 Hz for F1 and F2 respectively).

Table 4. 12 Descriptive statistics for the vowel /ʊ/

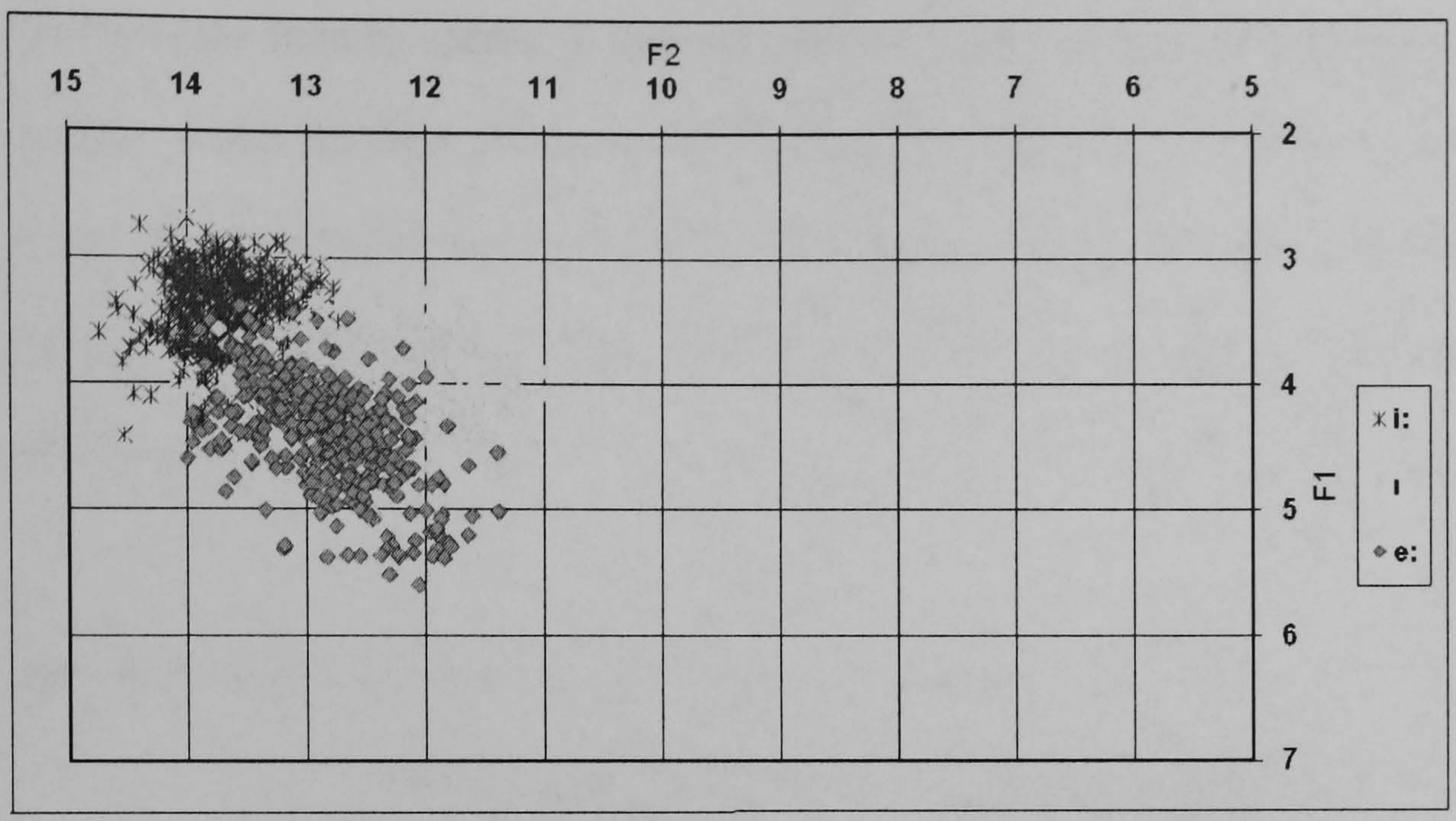
	Maximum	Minimum	Range	Mean	SD
F1	613	322	291	443	48
F2	1357	765	592	1026	133

The average formant values places this vowel somewhere in the middle between the highest and lowest average values reported for other Arabic dialects. The lowest F1 value was found in Iraqi (Al-Ani 1970), which is 290 Hz, and the highest value was reported by Abou Haidar (1994) for Jordanian at 580 Hz. On the other hand, the lowest F2 value was found in Iraqi (Al-Ani 1970) which is 800 Hz and the highest value was reported by Alghamdi (1998) for Sudanese at 1308 Hz.

Vowel /eː/

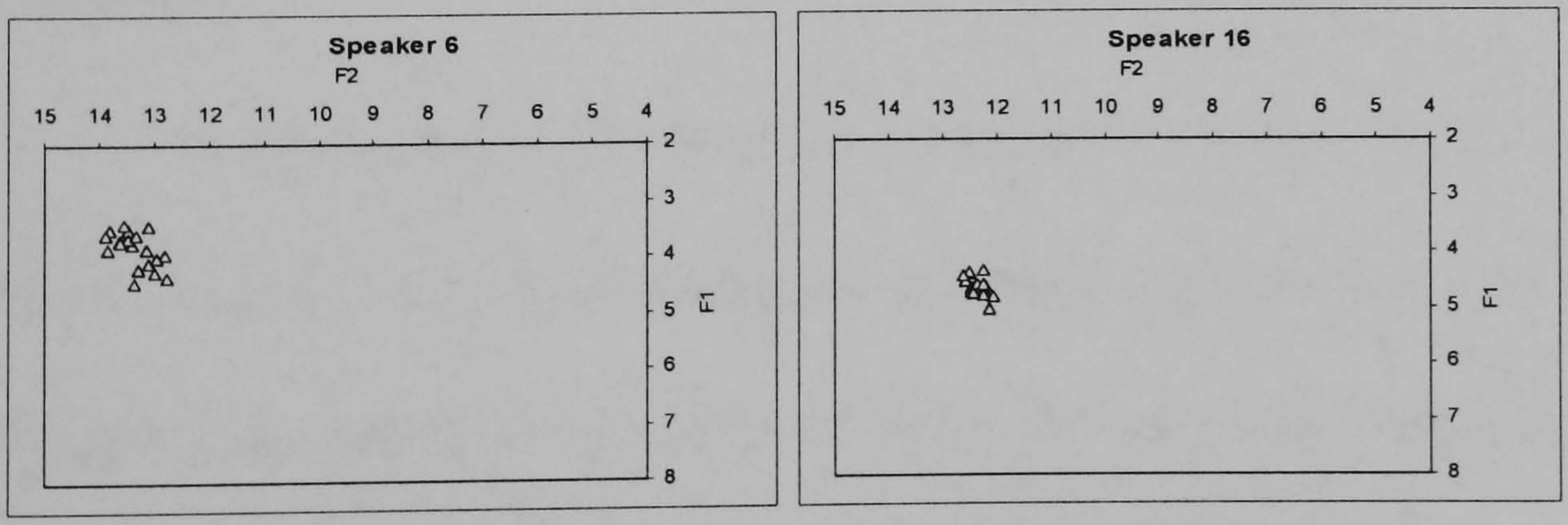
This is a high front long vowel as in /beːt/*(home)*, /deːn/ *(debt)*, and /zeːn/ *(beautiful)*. This vowel shares most of the acoustic area occupied by the high front short vowel /ɪ/ and some of the area occupied by the high front long vowel /iː/. However, /eː/ tends to be further back than /iː/ but more front than /ɪ/, and lower than both the latter vowels as the realisations of the three vowels by the 20 speakers in Figure 4.12 show.

.4. 12 Acoustic realisations of the vowels /iː/, /ɪ/ and /eː/ by the all speakers (Bark)



This vowel generally tends to show less variation within speakers compared to the previously discussed vowels. The realisations of the vowel tend to cluster closely around the mean and occupy a narrower space when produced by all speakers (see Appendix 2) apart from some exceptions. The productions of two speakers in Figure 4.13 are examples for comparison.

Fig.4. 13 Acoustic realisations of the vowel /e:/ by speakers 6 and 16 (Bark)



As can be seen from these representations, this vowel is realised higher and further to the front by speaker 6 than by speaker 16. This observation is supported by the formant means for these speakers. The F1 mean for speaker 6's realisation of this vowel is 379 Hz while the F2 mean is 2107 Hz, and the F1 mean for speaker 16's realisation is 468 Hz and the F2 mean is 1814. Table 4.13 shows some descriptive statistics for the vowel /e : /.

Table 4. 13 Descriptive statistics for F1 and F2 realisations of the vowel /e : / (Hz)

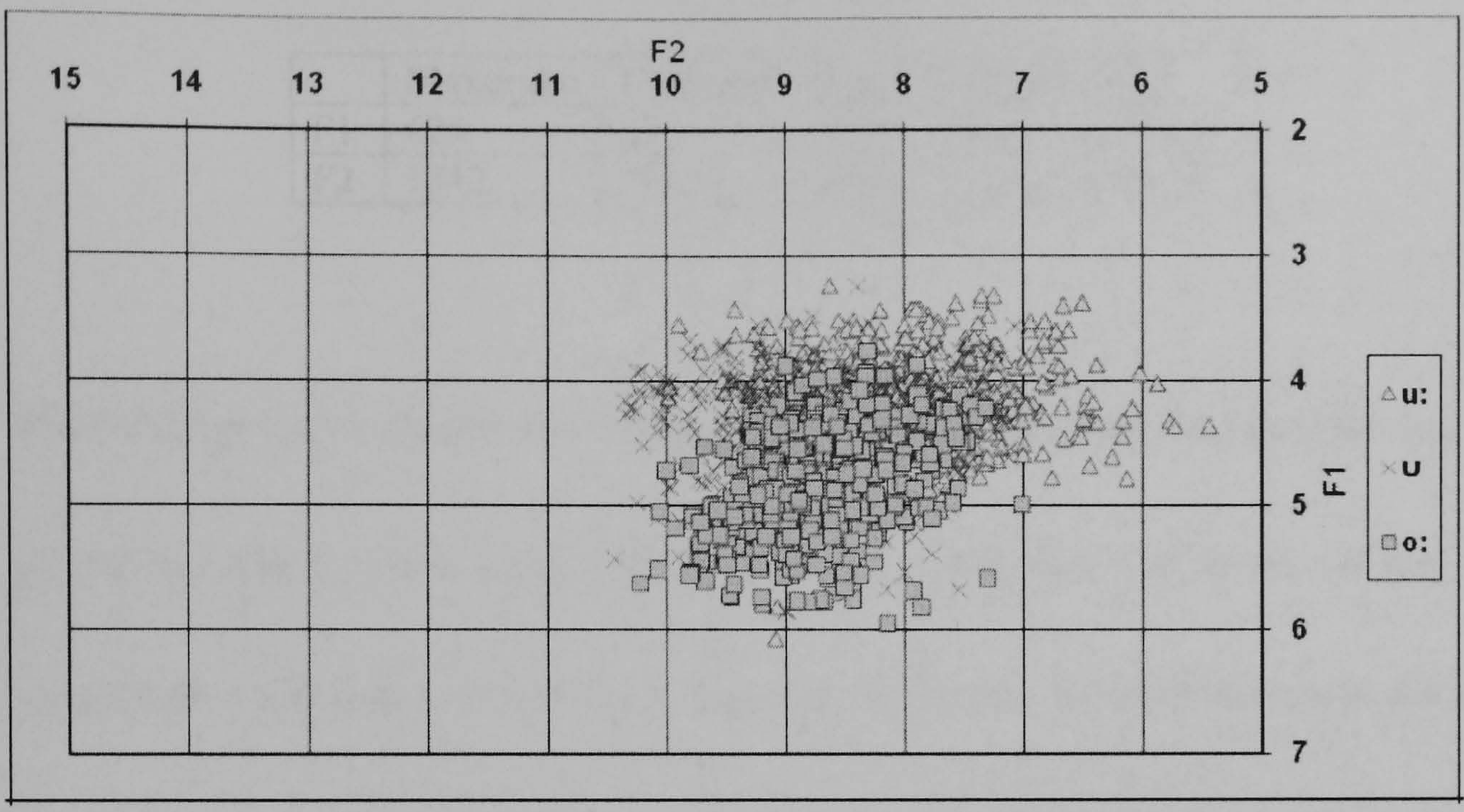
	Maximum	Minimum	Range	Mean	SD
F1	582	335	247	442	45.32
F2	2318	1568	750	1950	144.49

Although /e:/ is reported to be available in a considerable number of Arabic dialects, there are no reported formant values for this vowel. This might be attributed to the fact that it is not found in the standard variety of Arabic to which these dialects were compared.

Vowel /o : /

This is a mid back vowel as in /ko : n/ *(universe)*, /lo : n/ *(colour)* and /do : r/ *(level)*. The vowel /o : / overlaps considerably with the two high back vowels /u : / and /ʊ/ though it is generally lower and more to the front in the vowel space as shown in Figure 4.14 below. (A discussion of the vowels' overlap is given in section 4.2.1.3).

Fig.4. 14 Acoustic realisation of the vowels /u:/, /ʊ/, /o:/ by all speakers (Bark)



The distribution of the realisations of this vowel in the acoustic vowel space tends to be somewhat spread for all speakers (see Appendix 2) with the exception of speaker 2 whose tokens cluster together in the vowel space as Figure 4.15 shows.

There is less variation with regard to the F2 dimension than for the other two vowels with which this vowel overlaps, namely /u:/ and /ʊ/. However, /o:/ shows more variation with regard to the F1 dimension, as can be seen when the range and the SD for both formants in Table 4.14 are compared with those in Tables 4.12 and 4.13.

Fig.4. 15 Acoustic realisation of the vowel /o:/ by two speakers (Bark)

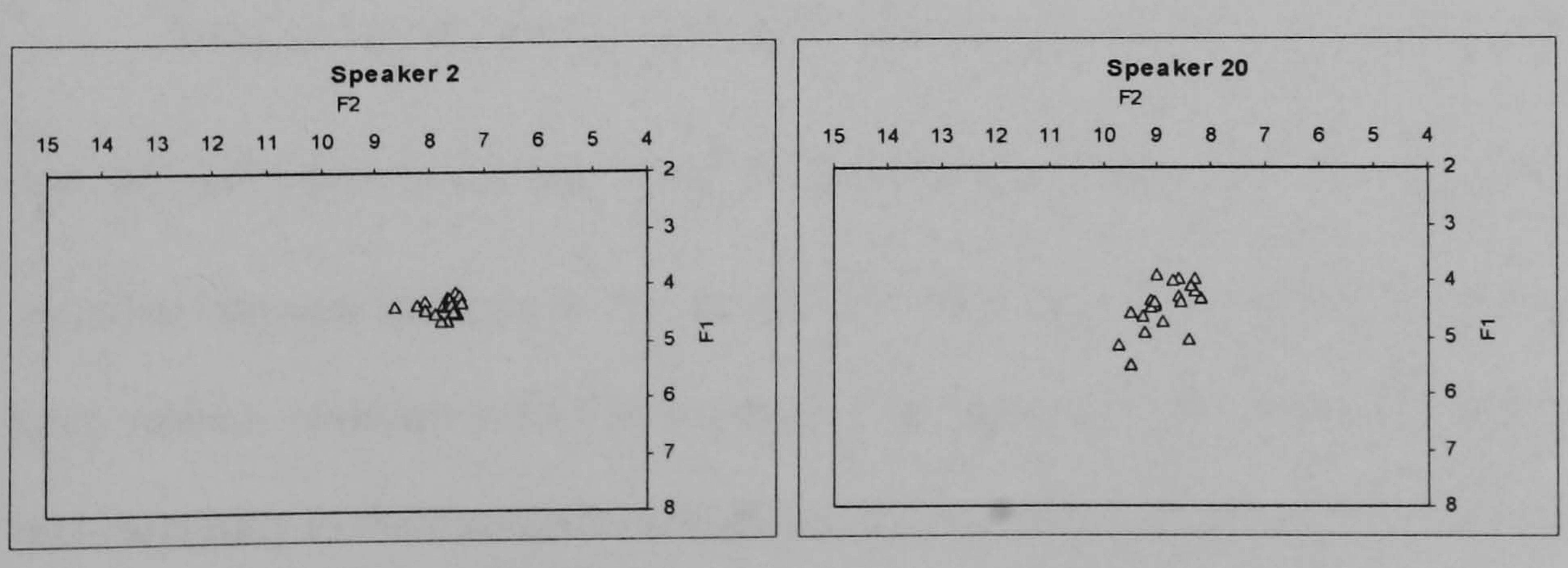


Table 4. 14 Descriptive statistics for the F1 and F2 realisations of the vowel /o : /

	Maximum	Minimum	Range	Mean	SD
F1	625	375	250	496	54.24
F2	1312	769	543	1037	96.33

The F1 range is 291 Hz and 321 Hz for /ʊ/ and /u : / respectively compared to only 250 Hz for /o:/. On the other hand, the F2 range is 592 Hz and 704 Hz for /ʊ/ and /u : / respectively compared to 543 for /o : /. The SDs for the two formants also show the same pattern, except for F1 where the vowel /o : / has the highest SD and the lowest range. The SDs for F1 are 47.79, 49.11 and 54.24 for /ʊ/, /u : / and /o : / respectively compared to those for F2 which are 133.30, 125.24 and 96.33 for /ʊ/, /u : / and /o : / respectively.

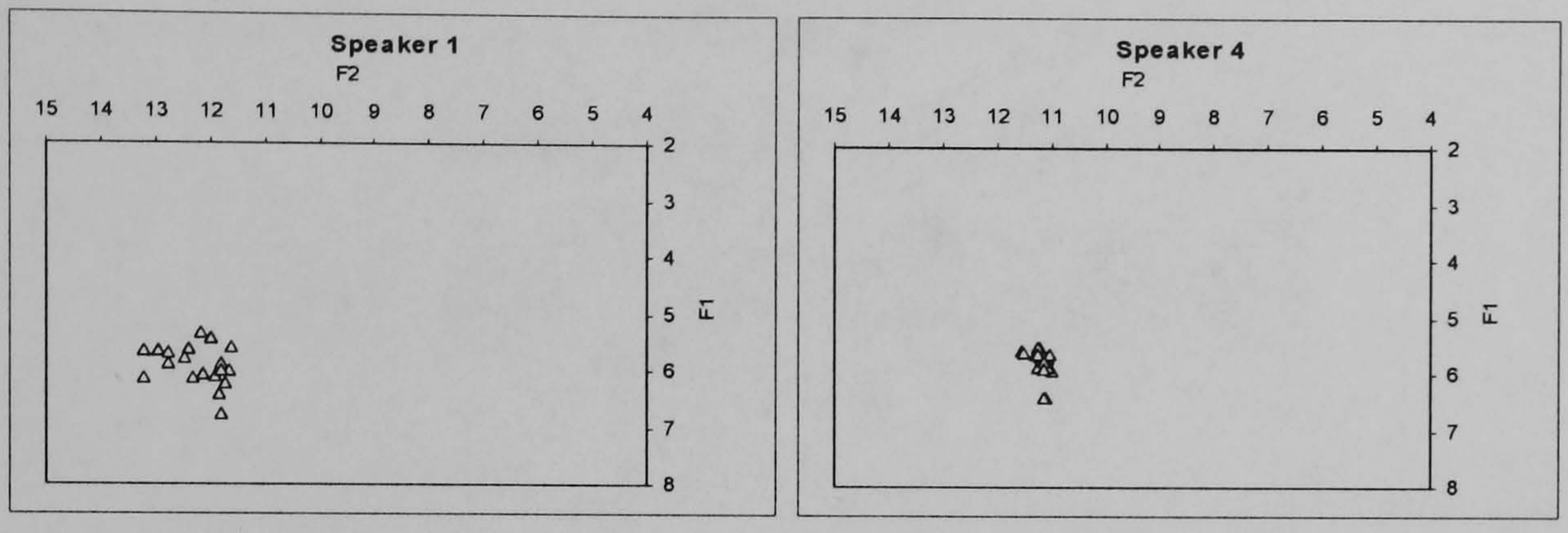
This mid back /o : / vowel, similarly to the mid front /e : /, does not have a short counterpart to which it can be compared. Thus, this vowel has not been considered by most of the acoustic studies, especially those comparing Arabic dialects to MSA where this vowel does not exist.

Vowel /æ : /

This is a low front long vowel as in /sæ : m/ *(poisonous)*, /gæ : s/ *(measured)*, and /læ : m/ *(blamed)*. As was found for the previously discussed vowels, there is some variation between speakers in the realisation of this vowel. While some speakers tend to have variable realisations of this vowel (e.g. speaker 1), others have a tendency to show less variability in their acoustic realisations (e.g. speaker 4) as shown in Figure 4.16.

Moreover, some speakers' productions tend to show more variation along the front-back dimension while others show more variation in vowel height. For instance, the

Fig.4. 16 Acoustic realisations of the vowel /æ:/ by speakers 1 and 4 (Bark)



F1 range for speaker 3 is 138 Hz compared to only 64 Hz for speaker 5. However, the F2 range for speaker 3 is 143 Hz compared to 392 Hz for speaker 5. Figure 4.17 shows the production by these two speakers of the vowel /æ:/.

Fig.4. 17 Acoustic realisations of the vowel /æ:/ by speakers 3 and 5 (Bark)

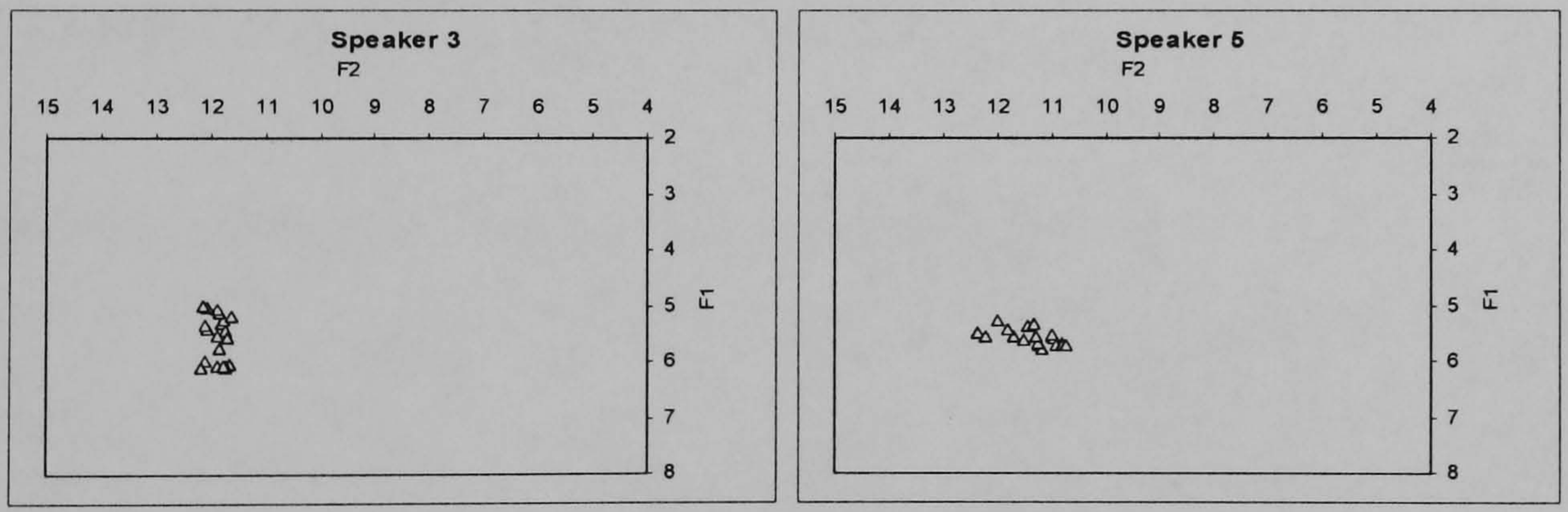


Table 4.15 includes some statistical results for the vowel /æ:/. The average formant values place this vowel somewhere in the middle between the highest and lowest average values reported for other Arabic dialects. The lowest F1 value was found in Tunisian (Belkaid 1984), which is 425 Hz, and the highest values was reported by Abou

Haidar (1994) for Jordanian at 770 Hz. On the other hand, the lowest F2 value was found in Iraqi (Al-Ani 1970), which is 1200 Hz, and the highest values was reported by Abou Haidar (1994) for Tunisian at 1780 Hz.

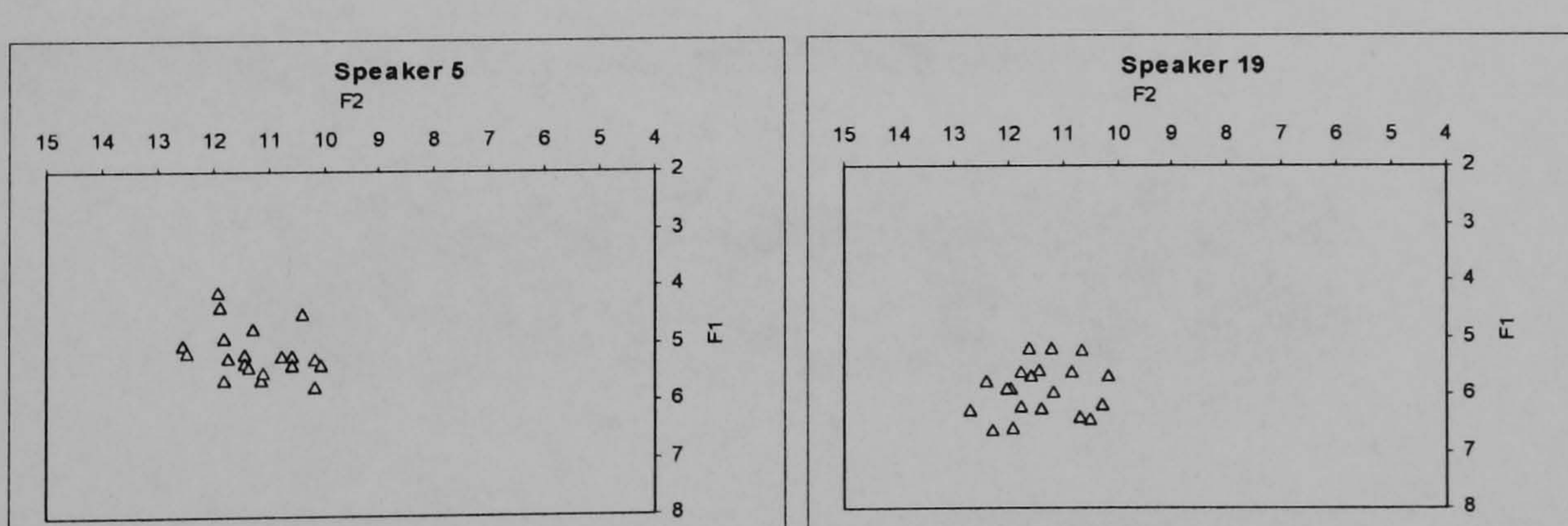
Table 4. 15 Descriptive statistics for the F1 and F2 realisation of the vowel /æ:/

	Maximum	Minimum	Range	Mean	SD
F1	733	456	277	588	45
F2	2155	1295	860	1641	139

Vowel /ə/

This is a low front short vowel as in /səb/ *(cursed),* /rən/ *(rang),* and /məs/ *(touched)*. A representation of the acoustic realisations of the vowel by the 20 speakers in the acoustic space shows that this short vowel has wide variation within and between most speakers. It also has a tendency to be central in the vowel space in the production of some speakers. Figure 4.18 shows some examples.

Fig.4. 18 Acoustic realisations of the vowel /ə/ by some speakers (Bark)



Compared to the low front long vowel /æ:/, with which the low front short vowel /ə/ considerably overlaps, vowel /ə/ tends to be higher and more centralised (see Figure 4.19).

The qualitative difference between the two vowels is also manifest in the descriptive statistics of their formant frequencies. Compare, for example, the formant means for /ə/ in Table 4.16 below to those for /æ:/ in Table 4.15 above.

Fig.4. 19 Acoustic realisations of the vowels /æ:/ and /ə/ (Bark)

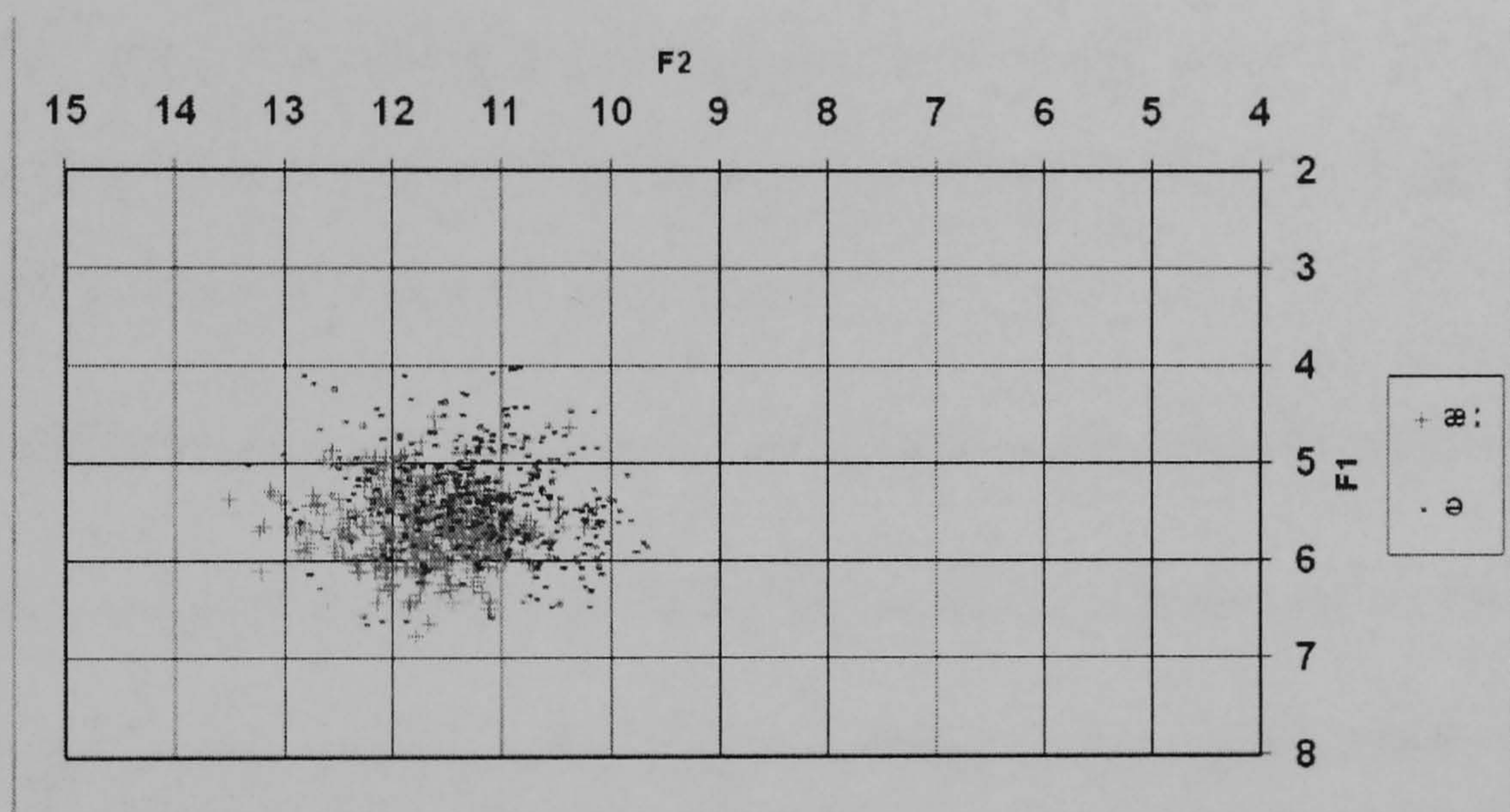


Table 4. 16 Descriptive statistics for F1 and F2 realisations of the vowel /ə/

	Maximum	Minimum	Range	Mean	SD
F1	717	401	316	555	60
F2	2105	1205	900	1541	165

When the F1 and F2 means of duration of these two vowels (see section 4.2.2) and the other vowels included in the LA vowel system were compared using a Two Sample (Kolmogorov-Smirnov) Test, it was found that the short vowel /ə/ and the other two short vowels /ɪ/ and /ʊ/ differ significantly from their long counterparts /æː/, /iː/ and /uː/ ($p < 0.001$) both qualitatively and quantitatively. As is already known, the tendency for this vowel to be more central than its long counterpart is shared with the other two short vowels in LA, the high front vowel /ɪ/ and the high back vowel /ʊ/.

The fact that the low short vowel /ə/ is higher than its long counterpart /æː/ has also been reported in other Arabic dialects, including Tunisian (Belkaid 1984), Iraqi (Al-Ani 1970), Syrian (Abou Haidar 1994), Saudi (Abou Haidar 1994; Alghamdi 1998), and Sudanese (Abou Haidar 1994; Alghamdi 1998). It should be admitted, however, that the difference is only 10 Hz for Syrian and Sudanese (Abou Haidar 1994). On the other hand, Qatari (Abou Haidar 1994) was reported to show almost no difference in height between the two low vowels /aː/ and /a/ (620 vs. 621).

However, there are some other Arabic dialects in which the low long vowel /aː/ (i.e. /æː/ in LA) was reported to be higher than its short counterpart /a/ (i.e. /ə/ in LA). These are Tunisian (Abou Haidar 1994), Jordanian (Abou Haidar 1994), Egyptian (Alghamdi 1998), Cairene (Newman 2002), Lebanese and Emirati (Abou Haidar). Again, it should be admitted, however, that the difference is only 10 Hz for Jordanian and for Emirati and only 6 Hz for Egyptian.

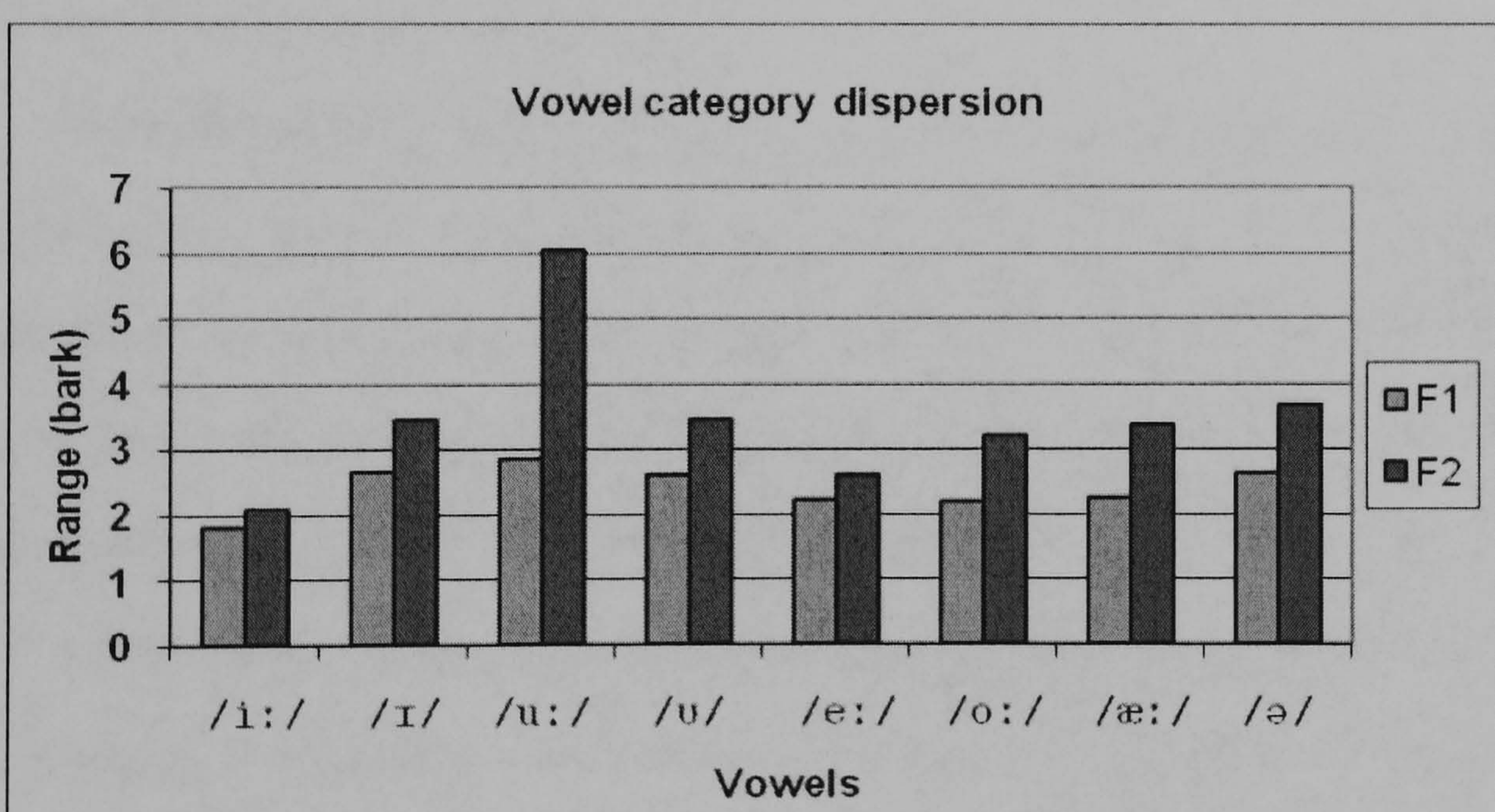
With regard to the front-back dimension, Arabic dialects vary according to which vowel is considered more front than the other. While results from some studies agree with those found in this study regarding the short vowel /a/ being more retracted, the same vowel has been shown to be further to the front than the low long vowel /aː/ in other dialects. The former group of dialects includes Tunisian (Abou Haidar 1994; Belkaid 1984), Saudi (Alghamdi 1998), Egyptian (Alghamdi 1998), Cairene (Newman 2002), and

Lebanese (Abou Haidar 1994). The latter group includes Iraqi (Al-Ani 1970), Syrian, Jordanian, Saudi, Qatari, Emirati (Abou Haidar 1994) and Sudanese (Abou Haidar 1994; Alghamdi 1998).

4.2.1.2. Between-vowel category variability

It has been noticed from the results above that variation occurred within as well as between vowel categories. In order to compare the within-category variability between all categories, the range of the dispersion[15] of each vowel category for both formants was calculated in bark. Then, this dispersion for all vowel categories was represented by a bar graph in Figure 4.20 to facilitate the comparison.

Fig.4. 20 Vowel category dispersion (Bark)



The long back vowel /u:/ appears to be the most dispersed vowel on the front-back dimension, having a range of more than 6 barks for F2. On the other hand, the least

15 The range of dispersion is the difference between minimum and maximum values of formants.

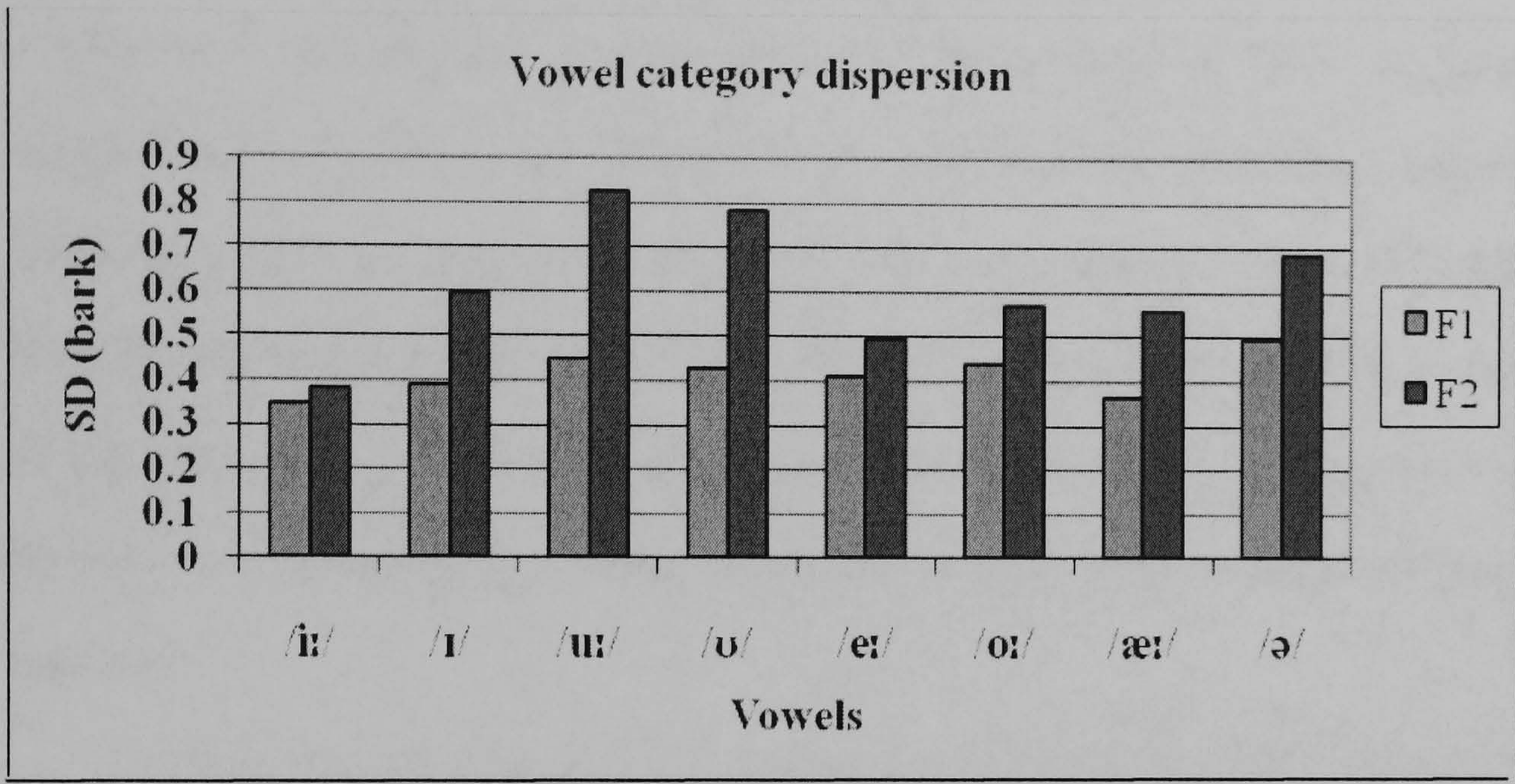
dispersed vowel seems to be the high front long vowel /iː/, with an F2 range difference of more than 2 barks. This vowel is also the least dispersed vowel with regard to F1. It has a range difference of less than 2 barks. All the other vowels show some kind of similarity in dispersion between them regarding F1 and F2.

Beckman, Jung, Jung, Lee, de Jong, Krishnamurthy, Ahalt & Cohen (1995) explained the fact that front vowels are less dispersed than back vowels by suggesting that the articulatory configuration of /i/ is easier to obtain than that of /u/, resulting in less variability in the formants. High front vowels can be produced more precisely by stiffening the genioglossus muscle and sustaining the tongue laterally against the dental edge. By contrast, the articulation of the back vowel /u/ cannot be obtained in a similar manner, leading to the less precise control of tongue height which results in more dispersion in the production of this vowel.

However, although range might give some idea about the dispersion of vowels, especially with regard to minimum and maximum points, it might not be the optimal measurement since there may exist some outliers that affect the measured dispersion in spite of the fact that they might be few in number. Another way to measure dispersion is to use standard deviation as shown in Figure 4.21 below.

Compared to Figure 4.20, Figure 4.21 does not show much difference in the overall pattern of dispersion. Here vowel /uː/ is still the most dispersed vowel, if F2 is considered, followed by the short vowel /ʊ/ which is slightly less dispersed. However, the vowel /ə/ is the most dispersed vowel as far as F1 is concerned. The least dispersion in both formants is once again shown by the long vowel /iː/. Apart from /ʊ/, it can generally be concluded that short vowels are more dispersed than long vowels and high front vowels are less dispersed than back and low vowels.

Fig.4. 21 Vowel category dispersion (SD)



A possible reason for this trend is the phonetic context in which these vowels occur (Frieda 2000, p.140). Unlike long vowels, short vowels are often undershot and therefore their formant steady state is never reached (Strange 1989, p.2082). The extent of this effect depends on several factors, including the type of consonants preceding and following the vowel and the speaking rate, among other factors (Lindblom 1963).

The tendency for high front vowels to be less dispersed than other vowels might be explained by the fact that most of the consonants preceding and/or following all vowels in the data material are coronals which have an anterior place of articulation which is similar to that used in the articulation of the high front vowels. The tongue does not need to move over a long distance to/from the consonant in order to produce the vowel. This leads to less influence on front rather than back vowels and makes them less dispersed. Moreover, low vowels require jaw lowering, which requires more time for the tongue to move to/from the articulatory position of a low vowel, which makes this vowel more liable to consonantal effects.

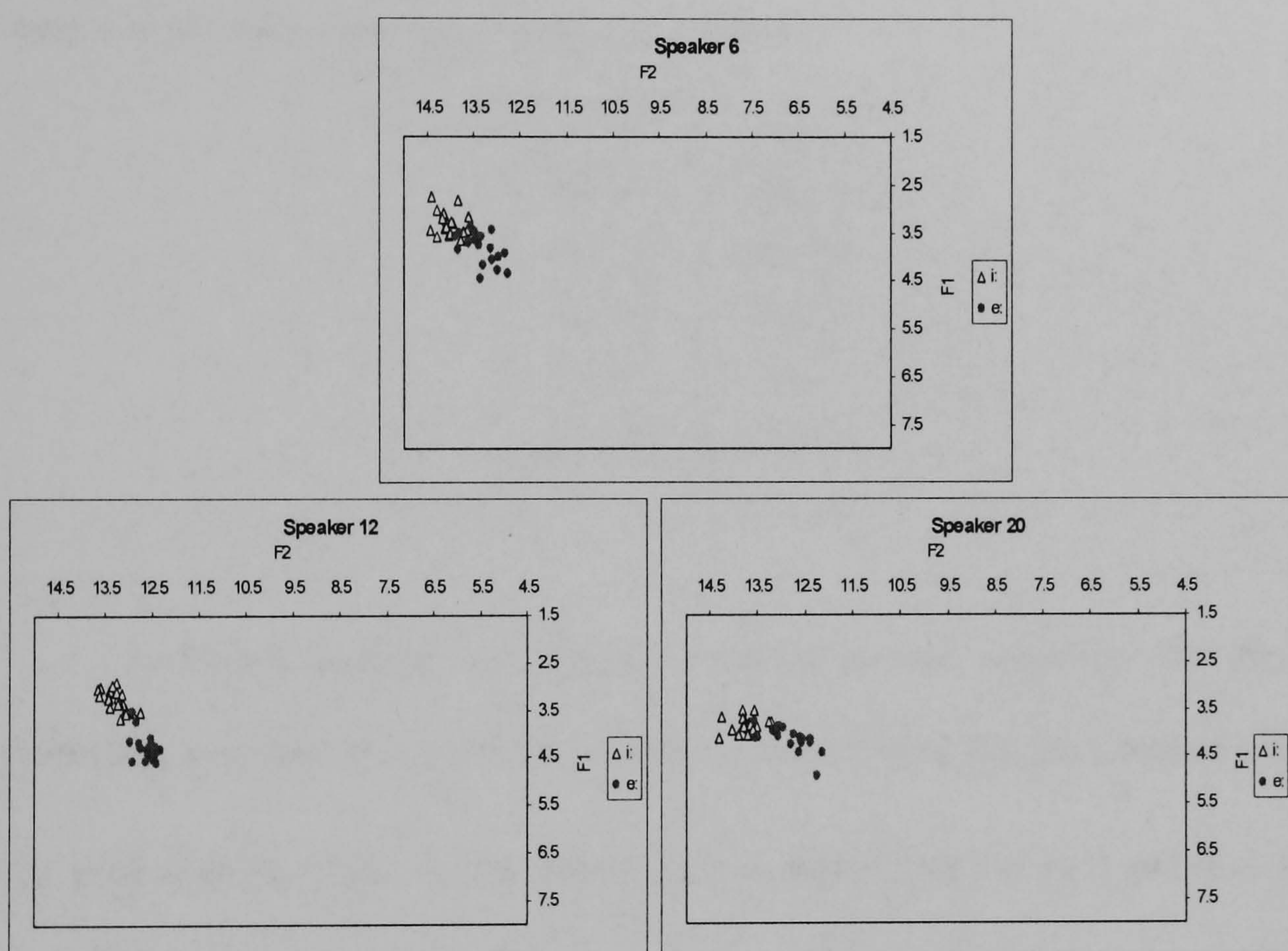
The effect of the distance between a vowel position and a neighbouring consonant position is well documented in the literature. For example, Hussain (1985) and Ghazeli (1977) found that emphatic consonants lead to the lowering of F2 in neighbouring vowels. The F2 of vowels is also affected by the duration of the vowel. Thus, long vowels tend to have higher F2 values than short ones in the same context (Hussain, 1975; Ghazeli 1977). Rosner and Pickering (1994, p.22) maintain that the degree of influence depends on the articulatory configuration of affecting and affected sounds. Therefore, vowel formants vary according to variation in the articulatory configuration of neighbouring consonants.

Finally, Quantal Theory claims that the fundamental vowels /i/, /u/ and /a/ are more stable in the vowel space and thus less dispersed than other vowels (Steven 1972). In LA this seems to be true of /i:/ and /æ:/, which show less dispersion than all other vowels. However, this tendency has not been found true for /u:/, which is more dispersed than the other vowels contained in the inventory including its short counterpart /ʊ/. The claim made by quantal theory that fundamental vowels are less dispersed than other vowels in the same inventory was also found to be true in Moroccan and Jordanian Arabic dialects (Al-Tamimi and Ferragne 2005, p.4).

4.2.1.3. Vowel quality overlap

When vowels were presented on a formant plot in Figure 4.2, some vowels were found to overlap. This might be partially attributed to the high degree of variability these vowels exhibit. For example, the acoustic space for the two long vowels /i:/ and /e:/ was found to overlap with that of the short vowel /ɪ/. Similarly, the two long vowels /u:/ and /o:/ were also found to overlap with the short vowel /ʊ/. However, a closer examination of the data shows that some speakers are responsible for most of the overlap. For instance, it was found that only three speakers exhibit overlap between the two vowels /i:/ and /e:/. Figure 4.22 shows these overlapping vowels as produced by these speakers.

Fig.4. 22 Acoustic realisations of /i:/ and /e:/ by some speakers showing overlap between the two vowels (Bark)



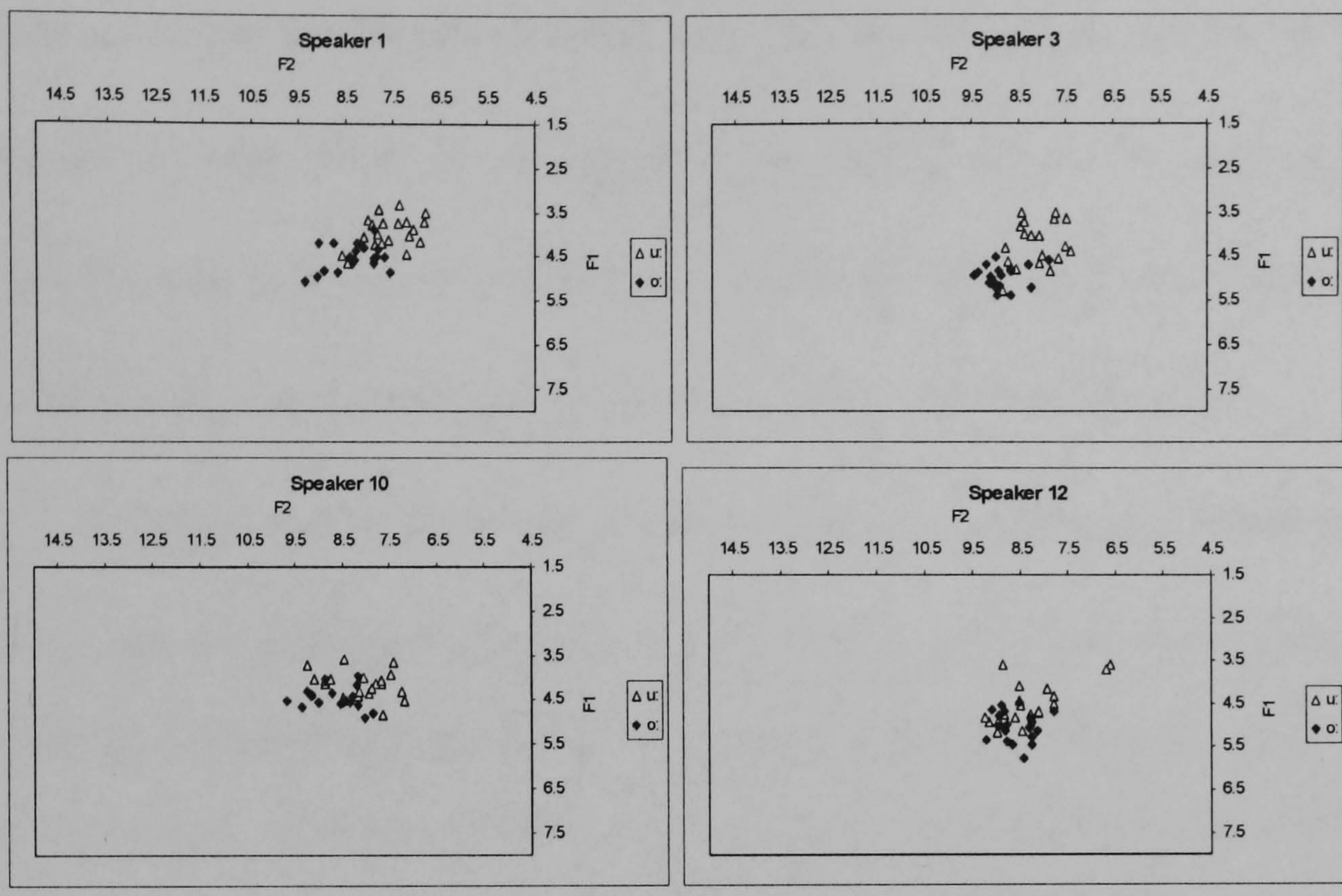
The words containing these overlapping vowels are shown in Table 4.17 below. When the researcher listened to these words, the vowels contained in them were found to be perceptually distinctive. What enhances these observations is that no minimal pairs which only differ by the vowel are found among the overlapping tokens in spite of their existence in the material tested, such as /di:n/-/de:n/. In other words, when the context in which these vowels occur is the same, there would be less possibility of overlapping between these vowels when they are produced by the same speaker. This might indicate that the different contexts in which these vowels occurred are responsible for this overlap (Strange et al 1983, p.695). That is, listeners would normally rely on differences in context for the identification of these vowels when the internal cues to the vowel are not sufficient.

Table 4. 17 Word tokens including overlapping /i:/ and /e:/

/i:/	/e:/
/bi:k/	/de:n/
/gi:s/	/ʒe:b/
/si:d/	/ze:d/
/si:n/	/ze:n/
/li:g/	/ke:l/

A similar acoustic and auditory examination was conducted for the words containing overlapping /u:/ and /o:/ formants. It was found that the extent of overlap in the case of these vowels in the vowel space is higher than for /i:/ and /e:/. Only 7 speakers from the twenty speakers participating in the experiment did not show overlap in these two vowels. The other 13 speakers showed varying degrees of overlap. Figure 4.23 shows some examples.

Fig.4. 23 Acoustic realisations of /u : / and /o : / for some speakers showing overlap between the two vowels (Bark)



In order to examine whether these vowels are acoustically and auditorily distinctive, minimal pairs containing these vowels were selected. These are shown in Table 4.18.

Table 4. 18 Minimal pairs containing /u : / and /o : / vowels

Vowels	Minimal pairs
/u : /-/o : /	/dɔ : m/-/du : m/
	/mɔ : t/-/mu : t/
	/rɔ : z/-/ru : z/

Regarding the minimal pair /du : m/-/do : m/, speaker 1 realised the word /du : m/ as [do : m] and /mu : t/ as /mo : t/. However, this was not the case with the minimal pair

/ru:z/-/ro:s/ which were auditorily distinctive. Similarly, speakers 2 and 10 realised /du:m/ as /do:m/ but the other minimal pairs were distinct. Finally, Speaker 12 also produced the word /du:m/ as [do:m] and /ru:z/ sounded like [ro:z]. However, the vowels in the the minimal pair /mu:t/-/mo:t/ sounded different. Other words containing /u:/ such as /ru:f/ also showed that this vowel was realised differently from /o:/.

An explanation for the overlap between /o:/ and /u:/ is offered by Beckman et al (1995), who maintain that, in contrast with high front vowels which involve precise articulatory configurations resulting from stiffening the tongue and propping it against the dental rim, back vowels are produced with less control of the tongue which results in overlap between /u:/ and /o:/. Consequently, this will affect perception, making the listener's ability to distinguish back vowels weaker than that of front vowels.

Acoustic overlap in the vowel space is common. For example, Peterson and Barny (1952) found a considerable overlap in the production of American vowels by native speakers. As stated before, this overlap can be attributed to several factors, including differences in the vocal tract size (Ryalls 1996, p.33) and the context in which these vowels are produced. When there is variation in the vowel produced by the same speaker, listeners use structural estimation (Nusbaum & Morin 1992, p.131). That is, they rely on the different cues found in the vowel produced to identify it. For example, listeners benefit from F0 and F3 in addition to F1 and F2 in order to recognise the vowel. However, when different vowels produced by the same speaker sound similar, listeners rely on a contextual tuning mechanism (Nusbaum & Morin 1992, p.131). In this kind of normalisation, listeners try to benefit from contextual information external to the vowel. Information relating to other sounds and utterances found in the vicinity of that vowel is used in addition to cues found in higher levels including lexical, grammatical and semantic cues.

4.2.2. Duration analysis results

In order to present and discuss the results for the duration analysis, the overall duration for all vowels is presented first before looking closely at the duration patterns of each vowel separately.

Table 4. 19 LA vowel duration means in milliseconds

LA vowels	/iː/	/ɪ/	/uː/	/ʊ/	/eː/	/oː/	/aː/	/ə/
Max	245	107	257	115	256	246	241	109
Min	53	22	81	29	80	83	80	26
Range	192	85	176	86	176	163	161	83
Mean	138	54	148	64	156	154	150	63
SD	35	15	32	16	36	32	34	15

An examination of Table 4.19 reveals that Libyan vowels can be divided in two groups as far as duration is concerned: short and long vowels. The shorts vowels are three in number. They are the two high vowels /ɪ/ and /ʊ/ and the low vowel /ə/. On the other hand, the long vowels are five in number. These include the two high vowels /iː/ and /uː/, the two mid vowels /eː/ and /oː/, and the low vowel /æː/. The mean duration for the three short vowels together is 60 ms while that for the five long vowels is 149 ms and the ratio between the two groups is 0.40. Table 4.19 also shows the durational similarity within these two groups. High front vowels have the shortest duration in both groups. The Two Sample Test (Kolmogorov-Smirnov) results showing statistically significant relationships in duration between all vowels are presented in Table 4.20.

As expected, the results show that the difference between short and long vowels is significant. Unexpectedly, there is a significant difference in duration between long and short high front vowels /iː/ and /ɪ/, on the one hand, and their long and short counterparts in the vocalic system on the other (shaded cells). A significant durational relationship is also found between the long vowels /uː/ and /eː/ ($p = 0.022$).

Table 4. 20 Significance relations in duration between vowel categories[16]
(Shaded cells refer to unexpected significance between a long vowel and its long counterparts or between a short vowel and its short counterparts).

	/ɪ/	/uː/	/ʊ/	/eː/	/oː/	/æː/	/ə/
/iː/	0.001	0.003	0.001	0.001	0.001	0.004	0.001
/ɪ/		0.001	0.001	0.001	0.001	0.001	0.001
/uː/			0.001	0.022	0.111	0.393	0.001
/ʊ/				0.001	0.001	0.001	0.627
/eː/					0.792	0.220	0.001
/oː/						0.465	0.001
/æː/							0.001

High vowels being shorter than low vowels is attributed to the extra time needed for lowering the jaw when low vowels are produced (Lindblom 1968; Lehiste 1970). On the other hand, high vowels being shorter than back vowels can be explained by the fact that most of the contexts in which these vowels occur contain consonants that have a + anterior feature which is the same feature as for high vowels (see Table 4.2). That is, the tongue does not need to move over such a long distance from/to a consonant produced at the front of the oral cavity when high front vowels are produced. This distance will be longer when back vowels are produced, which leads these vowels having longer duration.

16 Those values which are less than 0.05 indicate significant relationships between the two vowels while those higher than 0.05 indicate that this relationship is statistically insignificant.

Fig.4. 24 Duration of /iː/ and /ɪ/ in ms as produced by all speakers (Bark)

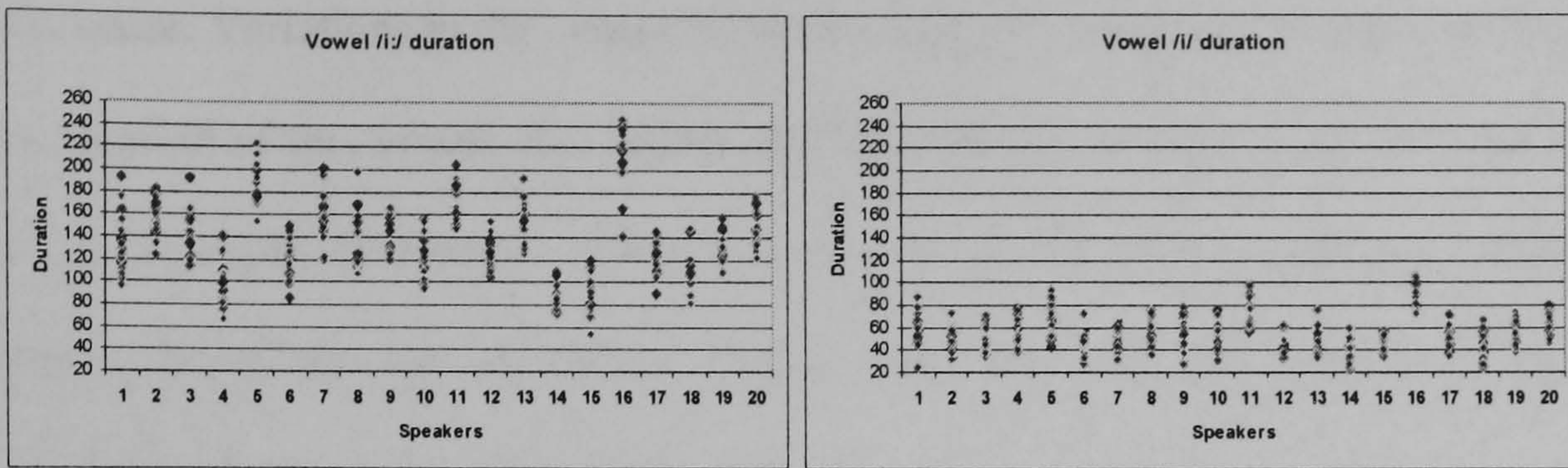
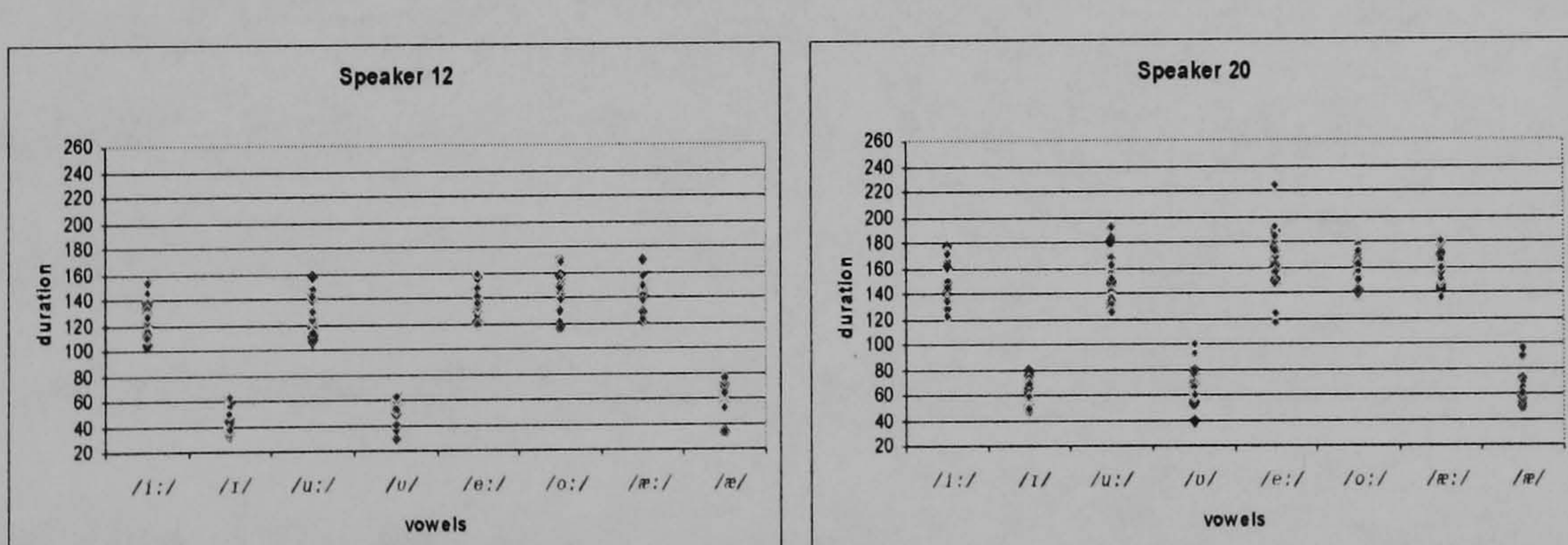


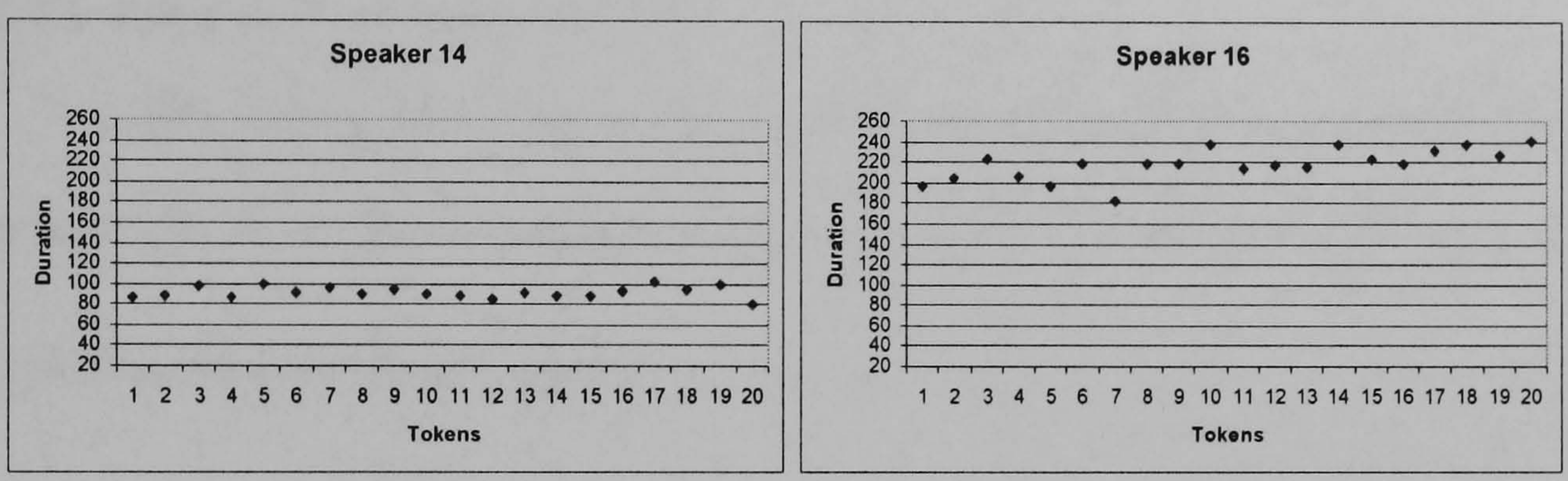
Figure 4.24 shows that there is considerable variation in durational patterns between speakers. However, the distinctive difference in duration between short vowels and long ones is still obvious, and when the duration of a short vowel is high, the duration of a long counterpart produced by the same speaker is still significantly higher in order to preserve the distinction in duration between short and long vowels. Compare, for example, the production by speakers 5 and 16 of the vowel /iː/ with their production of the vowel /ɪ/. The production of the other vowels follows the same overall pattern. Figure 4.25 shows the duration of all vowels in LA as produced by speakers 12 and 20.

Fig.4. 25 Duration of LA vowels in ms as produced by speakers 12 and 20 (Bark)



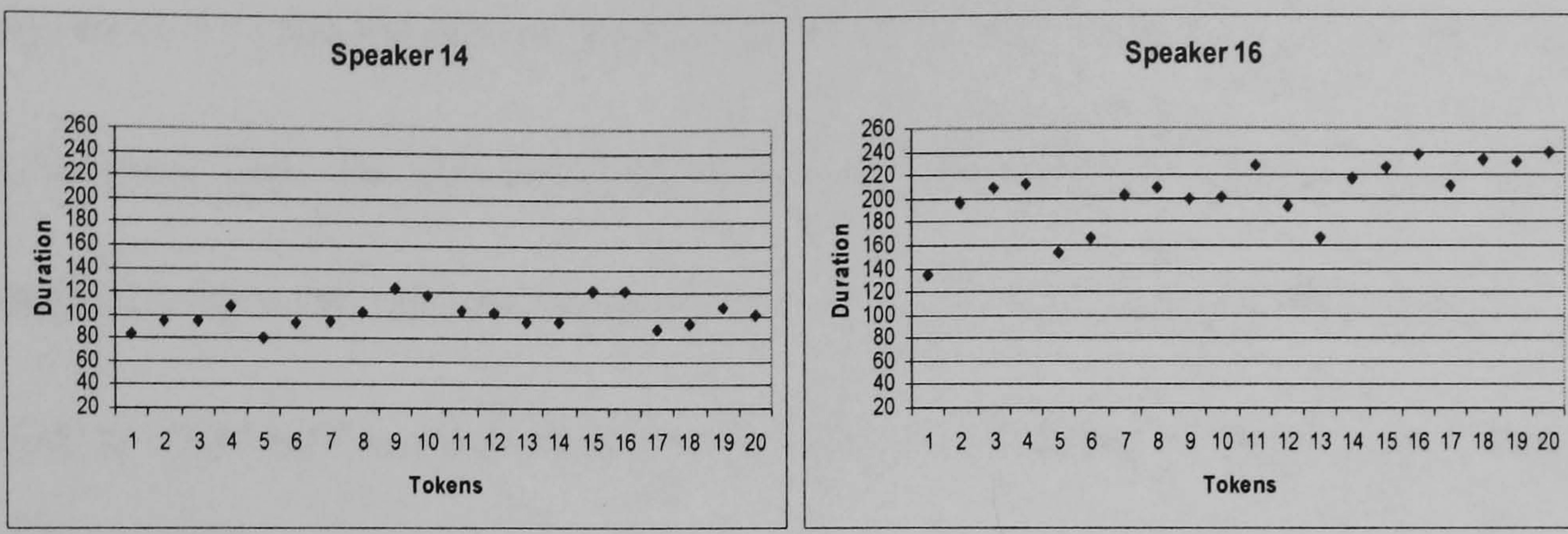
As can be noticed from these two representations, the durations of these vowels vary not only between vowels produced by the same speaker but also within each vowel produced by the same speaker. An apparent reason might be the context in which these vowels occur. Variations in the context in which a single vowel occurs result in variations in the duration of this vowel. But before dealing with the context in which these vowels occur, an example of a single vowel and how it is realised by different speakers in different contexts can be considered. Figure 4.26 represents the duration of the long vowel /æ:/ as produced by speakers 14 and 16.

Fig.4. 26 Duration of the vowel /æ:/ as produced by speakers 14 & 16 (Bark)



The variation of the duration of this vowel between the two speakers is apparent. While the maximum duration for speaker 14 is 102 ms, it is 241 ms for speaker 16. The range between the maximum production and the minimum production is 22 ms for speaker 14 and 59ms for speaker 16. The difference between the two speakers in duration might partially be attributed to the speaking rate at which the speakers produce the tokens (see Figure 4.27). In fact, a similar trend is found when looking at the realisations of the /u:/ vowel by the same speakers, except that the variation is greater within speakers in the production of /u:/ than of /æ:/. The difference in duration between high and low vowels will be considered in more detail in a separate section.

Fig.4. 27 Duration of the vowel /u:/ as produced by speakers 14 & 16 (Bark)



4.2.2.1. Long vs. short vowel duration

In this section, short vowels and their long counterparts will be compared. That is, the vowels /ɪ, ʊ, ə/ will be compared with /i:, u:, æ:/. First, some descriptive statistics of duration are introduced in Table 4.21.

Table 4. 21 Long vs. short vowel duration

	/i:/	/ɪ/	/u:/	/ʊ/	/æ:/	/ə/
Max	245	107	257	115	241	109
Min	53	22	81	29	80	26
Range	192	85	176	86	161	83
Mean	138	54	148	64	150	63
Ratio	0.39		0.43		0.42	
SD	35	16	32	16	34	15

As Table 4.21 shows, long vowels are more than twice as long as their short counterparts. The average mean for short vowels is 60ms and that for long vowels is 145ms, with a ratio of 0.41. Compared with the spectral difference between these vowels, as discussed in section 5.1.1, duration seems to represent a more robust distinguishing factor between short and long vowels. Table 4.21 also shows that the shortest duration is exhibited by the high front vowel group. For instance, the difference between the high long vowel /iː/ and the nearest duration which is found in the high back vowel is 10ms. On the other hand, the durational difference between the high short vowel /ɪ/ and the nearest duration in the group of short vowels, which is found in /ə/, is 9ms. However, the groups of short and long high back vowels and low vowels do not show much difference in their duration.

Vowel duration in LA can be compared to that found in other Arabic dialects. Table 4.22 presents some descriptive statistics of duration in some of these dialects, including those obtained in this study for LA.

Table 4. 22 Duration of long and short vowels in some Arabic dialects[17]
(Shaded areas indicate minimum and maximum values. Iraqi vowels are shaded to mark their exclusion from the comparison due to their excessive values).

Dialect	Short vowels	Long vowels	Ratio
Libyan (this study 2008)	60	145	0.41
Saudi (Alghamdi 1998)	119	232	0.51
Sudanese (Alghamdi 1998)	120	291	0.41
Egyptian (Alghamdi 1998)	110	275	0.40
Iraqi (Al-Ani 1970)[18]	300	600	0.50
Jordanian (Mitleb 1984)	83	128	0.65
Gulf Arabic (Hussain 1985)	95	170	0.56

While Iraqi vowel duration seems to be drastically longer than that of the other Arabic dialects (where the average duration for Iraqi short vowels exceeds that of long vowels of the other Arabic varieties), this radical duration can be attributed to the fact that the vowels measured were produced in isolation and not in context. Therefore, Iraqi will be excluded from the comparison of these dialects with respect to their vowel duration.

[17] All mean values are rounded to nearest full numbers.

[18] It should be mentioned here that Al-Ani used these measurements to describe MSA but the participants he used were Iraqis.

Generally, the lowest average duration of both short and long vowels is found in LA (60ms and 145ms) and the highest is found in Sudanese also for both short and long vowels (120ms and 291ms). On the other hand, the highest ratio is found in Jordanian (0.65) while the lowest is found in Egyptian (0.40).

4.2.2.2. High vs. low vowel duration

Lowering the jaw is known to have a positive effect on vowel duration (Lindblom 1967; Klatt 1976). Therefore, it is expected that low vowels would be longer than high vowels due to the amount of jaw lowering required by the production of low vowels. This difference is evident in LA whereby /iː/ is shorter than /æː/, but the difference is only 2ms between /æː/ and /uː/. Moreover, the picture is not similar in the case of short vowels. Although the short low vowel /ɘ/ is longer than its high front counterpart (63 vs. 54 ms), this vowel has almost the same duration as that of its high back counterpart (63 vs. 64 ms). Moreover, the difference in length between the vowel /ɘ/ and its high front counterpart was found to be significant (see Table 4.18 above).

The finding that the low short vowel is unexpectedly slightly shorter than its high back counterpart can partially be attributed to the fact that the former has a similar realisation to the IPA /ɘ/ by some LA speakers. Another reason might be the context in which these two vowels occurred. As can be seen from Table 4.23, most of the consonants preceding and following these two vowels are anterior consonants. Moving the tongue from a consonantal anterior position to produce a back vowel like /ʊ/ requires more time than that needed for producing a low central vowel like /ɘ/ in the same context.

Table 4. 23 Words containing /ʊ/, /ə/ produced by all speakers

Vowels	Words									
/ʊ/	/buk/	/bul/	/ful/	/bud/	/bug/	/bum/	/bun/	/dub/	/hum/	/kub/
/ə/	/ləm/	/səd/	/fəd/	/nəg/	/səb/	/fək/	/rən/	/ħəm/	/məs/	/wən/

The fact that the degree of lowering of the jaw has a positive effect on the duration of the vowels is more evident when we compare high long vowels to mid long ones, as shown in Table 4.24.

Table 4. 24 Duration of high vs. mid long vowels

	/iː/	/uː/	/eː/	/oː/
Max	245	257	256	246
Min	53	81	80	83
Range difference	192	176	176	163
Mean	138	148	156	154
SD	35.31	32.03	35.75	31.62

The two high long vowels /iː and uː/ have average mean durations of 138ms and 148ms respectively, while the mid vowels /eː and oː/ have mean durations of 156ms and 154ms respectively. However, the latter two have longer durations than the low long vowel /æː/. This can partially be explained by the effects of the contexts in which these vowels occurred. For example, the low vowel /æː/ was followed by voiceless consonants in 120 tokens while the high vowel /eː/ was followed by voiceless consonants in only 80 tokens. The number of tokens in which these vowels are followed by voiced consonants is

320 tokens for the /e:/ and 280 tokens for the vowel /æ:/. The effect of context on vowel duration is important and, therefore, is dealt with in more detail in the next section.

The fact that low vowels are longer than high vowels is also manifest in other Arabic dialects. Table 4.25 compares the average means of low vowels to those of high vowels.

Table 4. 25 Average means of duration of high and low vowels in some Arabic dialects

Dialect	a (ə)	i (ɪ)	u (ʊ)	a: (æ:)	i:	u:
Libyan (this study)	63	54	64	150	138	148
Saudi (Alghamdi)	133	111	114	311	248	137
Sudanese (Alghamdi)	128	117	116	295	275	304
Egyptian (Alghamdi)	122	98	111	316	255	253
Iraqi (Al-Ani)	300	300	300	600	600	600
Jordanian (Mitleb)	90	76	83	145	116	124
Gulf Arabic (Hussain)	106	85	93	190	155	165

Apart from Iraqi which shows the same values of duration for low vowels and their high counterparts, all other varieties show a tendency towards low vowels being longer than high ones. The only two exceptions are found in LA and Sudanese. In LA, the short vowel /ə/ has a lower duration than its high back counterpart, for the reasons mentioned above, and in Sudanese the back long vowel /u:/ has a slightly longer duration than its low counterpart.

4.2.2.3 The effect of consonant voicing on vowel duration

The effect of consonant voicing on vowel duration in Arabic has been studied by a small number of researchers (see section 3.1.2.1 in Chapter Three). However, the results are not consistent. While some researchers maintain that this effect is significant (for example, Port et al 1980, Hassan 1981), others have found this effect not to be significant (such as Flege 1979, Hussain 1985, Mitleb 1984a). In dealing with the effect of context on duration in LA, the analysis focuses on the effect of the voicing of the consonants preceding and following the vowel.

4.2.2.3.1. The effect of the voicing of the following consonant

Because the data collected did not contain a sufficient number of comparable minimal pairs, a different procedure was followed to investigate the effect of voicing of the following consonant. In this procedure, all words that ended in voiced consonants were grouped together separately from those which ended in voiceless consonants. Then, the mean for the duration of each vowel was calculated. Table 4.26 shows the results.

Table 4. 26 Duration of vowels before voiced and voiceless consonants

Vowels	/i:/	/ɪ/	/u:/	/ʊ/	/e:/	/o:/	/a:/	/ə/
Before vd	141	56	149	66	160	157	150	63
Before vls	130	46	145	56	139	148	150	58
Difference	11	10	4	10	21	9	0	5

Apart from the long vowel /a:/ in which the mean duration is the same for both categories, all other results show that the mean duration of the vowel is longer when it is followed by a voiced consonant than when it is followed by a voiceless one. Non-parametric Tests (Mann-Whitney U and Wilcoxon) were conducted to compare the duration of vowels followed by voiced consonants to those followed by voiceless ones; it was found that there is a significant difference between the means of the two groups. Table 4.27 shows some results of these tests.

Table 4. 27 Non-parametric test (Mann-Whitney U and Wilcoxon) results of the voice effect of the following consonant on vowel duration

Vowels	Before	Mean	SD	Sig. (p-value)
Short vowels	Vd consonant	62	17.53	0.001
	Vls consonant	54	14.00	
Long vowels	Vd consonant	152	32.58	0.001
	Vls consonant	143	33.72	

These results show that the difference between the average mean duration of short vowels before voiced and voiceless consonants is 8ms and that between long vowels before voiced and voiceless consonants is 9ms. However, it should be admitted that the number of words from which these results were derived which ended in voiceless consonants was smaller than that of words ending in voiced ones. Table 4.28 shows the number of each category for each vowel.

Table 4. 28 Number of tokens of vowels preceding voiced vs. voiceless consonants

Before:	/iː/	/ɪ/	/uː/	/ʊ/	/eː/	/oː/	/aː/	/ə/
Voiced	280	320	200	360	320	280	280	320
Voiceless	120	80	200	40	80	120	120	80
Total	400	400	400	400	400	400	400	400

The finding that vowels preceding voiced consonants are longer than those preceding voiceless ones can be enhanced by comparing the only relevant minimal pair found in the data. This minimal pair is /bʊk/-/bʊg/, the two words of which end in voiced and voiceless velar stops. These two words were compared using a 2 Related Samples Test (Wilcoxon), the results of which are presented in Table 4.29.

Table 4. 29 Wilcoxon 2 Related Samples Test results

Words	Mean	N. of tokens	SD	P-value
/buk/	56.75	40	13.237	0.048
/bug/	60.30	40	16.544	

These results show that the vowel in /bug/ is longer than that in /buk/ by only 3.55ms on average. They also reveal that participants were more variable when producing /bug/ than/buk/. This increase was found to be significant. The probability value was 0.048, a P-value smaller than the threshold value of 0.05 which indicates that there was a significant difference, though marginal, between the means of the two words.

The difference in the significance levels of the effect of the voicing of the following consonant on vowel duration obtained from all words together (i.e., $p < 0.001$) and that obtained from the minimal pair /bug/ - /buk/ (i.e., $p = 0.048$) indicates that this effect might not only be dependent on whether the consonant is voiced or voiceless but also on other factors including their place and manner of articulation; something which cannot be tested using the data in hand due to lack of sufficient minimal pairs from different consonantal categories. In fact, Hussain (1985) confirmed the effect of place of articulation of consonants on vowel duration. He found that there was a tendency for vowel duration to increase with the transition time required for the articulator to move from the vowel position to the consonant position (Hussain 1985, p.151).

From a cross-dialect perspective, Arabic dialects vary with respect to the effect of voicing of the following consonant. For example, Alghamdi (1990) found that vowels followed by voiced consonants are significantly longer than those followed by voiceless consonants in Saudi Arabic. The ratio between the two in monosyllabic words was 0.85.

Flege (1979) also found a small but non-significant effect in the same dialect. According to Flege, Arabic has a phonemic length contrast for both vowels and consonants which should be maintained by speakers (Flege 1979, p.64). This contrast might reduce the durational differences normally found before voiced/voiceless consonants.

Another Arabic dialect which did not show a significant effect of the voicing of the following consonants is Jordanian, contrary to what was found in English when the productions of speakers from the two languages were compared (Mitleb 1984a). Data produced by speakers of English showed that, unlike Arabic, English has a significant difference in vowel duration between vowels followed by voiceless and voiced consonants.

4.2.2.3.2. The effect of voicing of the preceding consonant

The effect of the voicing of the preceding consonant has received less interest from researchers. However, Flege (1979, p.67), for example, found some evidence of temporal compensation between the vowel and the preceding consonant in the initial position of a word. It was found that vowels are significantly longer after voiced consonants than after voiceless ones. Because an initial /s/, for example, is longer than an initial /z/, vowel duration is shortened after /s/ to compensate for this length. However, Flege admits that this effect needs further investigation. The data from the present study contain two relevant minimal pairs, each of which only differs in the voice of the first consonant, which may provide support for the existence of an effect of the voicing of the preceding consonant on vowel duration. These two minimal pairs are /di:n/-/ti:n/ and /gi:s/ -/ki:s/, the average durational means of which are shown in Table 4.30 below.

Table 4. 30 Duration statistics of the vowel /i:/ preceded by voiced vs. voiceless consonants

words	mean	N. of tokens	SD	P-value
/di:n/	144	40	36.66	0.001
/ti:n/	131	40	33.68	
/gi:s/	133	40	34.87	0.279
/ki:s/	131	40	37.52	

The Non-parametric Two Related Samples Test (Wilcoxon) results show that the voice of the preceding vowel does have some effect on vowel duration in the sense that vowels preceded by voiced consonants are longer than those preceded by voiceless ones in spite of the fact that the difference in the second minimal pair is only 2ms. These findings are consistent with Flege's (1979) regarding the effect of the preceding consonant on vowel duration. However, while this increase was found to be significant in the minimal pair /di:n/-/ti:n/ ($p < 0.001$), it was not in /gi:s/-/ki:s/ ($p = 0.279$). This might indicate that the degree of effect of the voice of the preceding consonant on vowel duration also depends on the articulatory configuration of the consonant itself. It appears here that the alveolar /d/ has more influence on vowel duration than the velar /g/. It should be emphasised, however, that these findings are only preliminary because of the small number of words investigated.

In sum, the discussion of durational cues shows that, compared with the spectral differences between vowels as discussed in section 3.1.1, duration seems to represent a robust distinguishing factor between short and long vowels. Nevertheless, these findings concerning the production of LA vowels need to be confirmed by those for the perception of the same vowels. The relevant data are presented in the next chapter.

4.3. Summary

In order to describe the LA vocalic system, a production task was designed which involved the use of a robust sample of speakers who produced all eight LA vowels in monosyllabic words embedded in carrier sentences. Each speaker produced 160 tokens for all eight vowels and the total number of tokens for the 20 speakers was 3200. Formant frequencies and vowel durations were measured. Measurements were double-checked to ensure reliability.

The research results show that the Libyan vowel system includes 8 vowels. These vowels are grouped in two categories: long /iː, uː, eː, oː, æː/and short /ɪ, ʊ, ə/. The results also show that short vowels are significantly qualitatively different from their long counterparts in that short vowels tend to be more central in the vowel space. This significant difference justifies the use of different symbols for short vowels and their long counterparts.

The realisations of LA vowels by the participants show a degree of variability in the vowel space which resulted in vowel quality overlap. This overlap was more evident in the back high vowels when realisations of these vowels for speakers separately were examined. This was attributed to the fact that high back vowels do not have precise articulatory configurations.

With regard to vowel duration, high front vowels have shorter duration compared to high back vowels, which was explained by the context in which these vowels occurred. Most of the elicited words contained consonants produced in the front of the oral cavity where high front vowels are produced. Producing high back vowels in the same context requires the tongue to move over a longer distance and thus more time is needed. High vowels were generally shorter than low vowels because producing low vowels requires the lowering of the jaw which consumes more time. As for the effect of consonant voicing on vowel duration, vowels followed by voiced consonants were significantly longer than those followed by voiceless one.

CHAPTER V

PERCEPTION EXPERIMENT

5.0. Introduction

Based on the results of the production task, three tests were conducted in the perception part of the study. First, it was observed from the results obtained from the production of LA vowels by native speakers that there is a considerable overlap between these vowels. So, one of the aims of the perception experiment was to see whether this overlap also exists in perception and, if so, what makes these overlapping vowels distinct. The first identification experiment was created to test this (see section 4.5.2). Second, it was also found that short vowels are not only different from their long counterparts in quantity but also in quality. The aim of the second and third tests in perception was to discover whether or not listeners also rely on spectral cues in addition to their reliance on durational cues in perceiving these vowels. That is, whether the quality of these vowels is important in addition to their quantity when listeners perceive long vowels and their short counterparts. Therefore, two other identification test were created (see section 5.5.2).

5.1. Design of the perception task

5.1.1. Rationale for choosing /iː, ɪ /

The data obtained from the production task was so extensive that it was necessary to limit the perception task to one pair of vowels. Neither space nor time would allow for the study of perception for all vowels. While the auditory qualitative difference between the two elements of the pair /uː, ʊ/ and those of the pair /aː, ɐ/ is also auditorily noticeable, the auditory analysis of /iː, ɪ / resulted in some of the most noticeable differences in vowel quality between /iː/ and / ɪ / in spite of the overlap between the two vowels in the production task (Appendix 5). This perception experiment aims at confirming these auditory analysis findings. Another reason is that this pair of vowels has been dealt with in other languages (e.g. Mora & Fullana 2007, Escudero 2000, Flege et al 1997) but has never been examined in LA; it is worth investigating it not only to benefit from the methods followed in other studies but also to compare listeners' response to varying acoustic cues for comparable vowel contrasts from a cross-linguistic point of view.

5.1.2. Stimuli

The research question related to the perception of LA vowels by native speakers requires the availability of vowel stimuli with acoustic characteristics which vary in controlled incremental steps. This was obtained by synthesizing natural vowels to create continua that are presented in a /C_C/ context to listeners participating in the study. Speech synthesis was used because it has been identified as an important tool for testing the importance of various acoustic cues which contribute to phonological distinctions (Fant, 1973, p.19).

In order to test listeners' reactions to durational and spectral differences in the short-long phonological contrast in Arabic, the /iː-ɪ/ pair was chosen due to the maximal spectral differences that were found in the realisation of these vowels by the twenty speakers. While the original aim was to use tokens from the production data for synthesis, the need for high quality laboratory data required new laboratory-based recordings by a native Libyan speaker (the author). After some experimentation with different phonetic contexts, a minimal pair was chosen whereby the initial and the final consonant are known to have a relatively stable locus in both perceptual and acoustic senses in order to allow for the creation of more natural-sounding synthetic stimuli. The preferred choice would have been a symmetric /d_d/ context because of the robustness of the /d/-locus, but the absence of such a word from the LA vocabulary led to choosing the /d_b/ context.

Several repetitions of the tokens /diːb/ 'wolf' and /dɪb/ 'walk slowly' were recorded directly into a PC with high quality sound card in a sound-proof room using Cool Edit software and a Shure SM58 dynamic microphone. The durational and spectral values for the vowels in the recorded tokens fell within the range that was found in the production task (Table 5.1). The tokens were therefore considered representative of the realisations of these vowels in LA.

Table 5. 1 Mean duration and formant frequencies (SD in brackets) for three tokens of /diːb/ *'wolf'* and /dɪb/ *'walk slowly'* as produced by a native Libyan speaker

/diːb/ duration = 238 (10)				/dɪb/ duration = 63 (1)			
	onset	middle	offset		onset	middle	offset
F0	165 (1)	190 (0)	140 (5)	F0	175 (3)	184 (5)	183 (5)
F1	347 (13)	317 (1)	320 (1)	F1	383 (8)	383 (8)	392 (15)
F2	2090 (21)	2307 (7)	2260 (16)	F2	1782 (13)	1830 (3)	1790 (32)
F3	2658 (51)	2732 (17)	2567 (5)	F3	2541 (23)	2561 (30)	2480 (42)

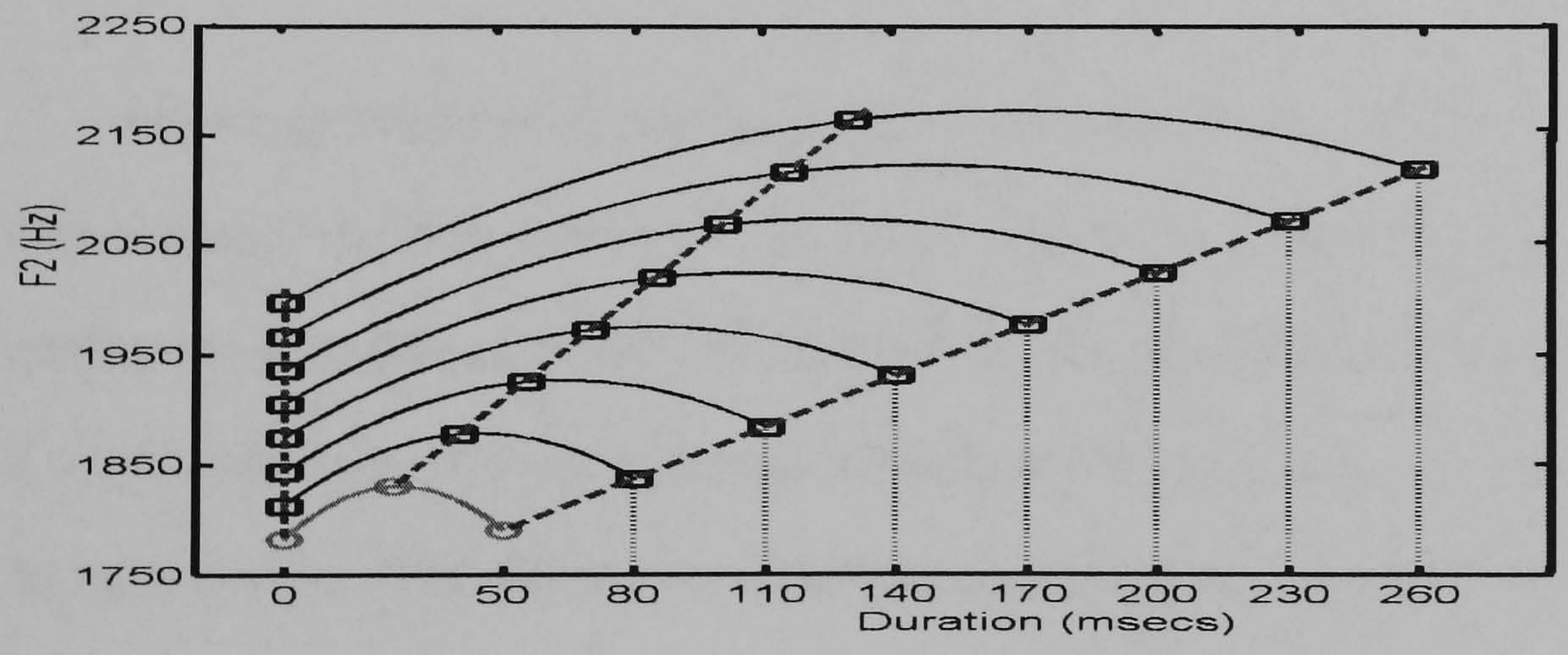
The spectral values obtained for the mid-point of the vowels and their transitions allowed the linguist[19] who helped with the synthesis to capture the main coarticulation

[19] Many thanks to Frantz Clermont for helping with the synthesis and creating the continua for the perception experiment.

effects through the syllable using a parametric model (Bergem 1994; 1995). This allowed him to estimate the dynamic structure of the formant transitions within the vowel nuclei and to create continua whereby short /dɪb/ tokens could still have dynamic spectral values that emulated the ones found in /diːb/, and long /d iːb/ tokens could still have spectral values that emulated the ones found in /dɪb/. The same estimation model was used to create the tokens for the remaining points along the continua. This required adapting the parabolas for the long and short formant shapes to the required durations using linear interpolation and parabolic morphing (Clermont, Harrison, and French, 2007) (see Figure 5.1).

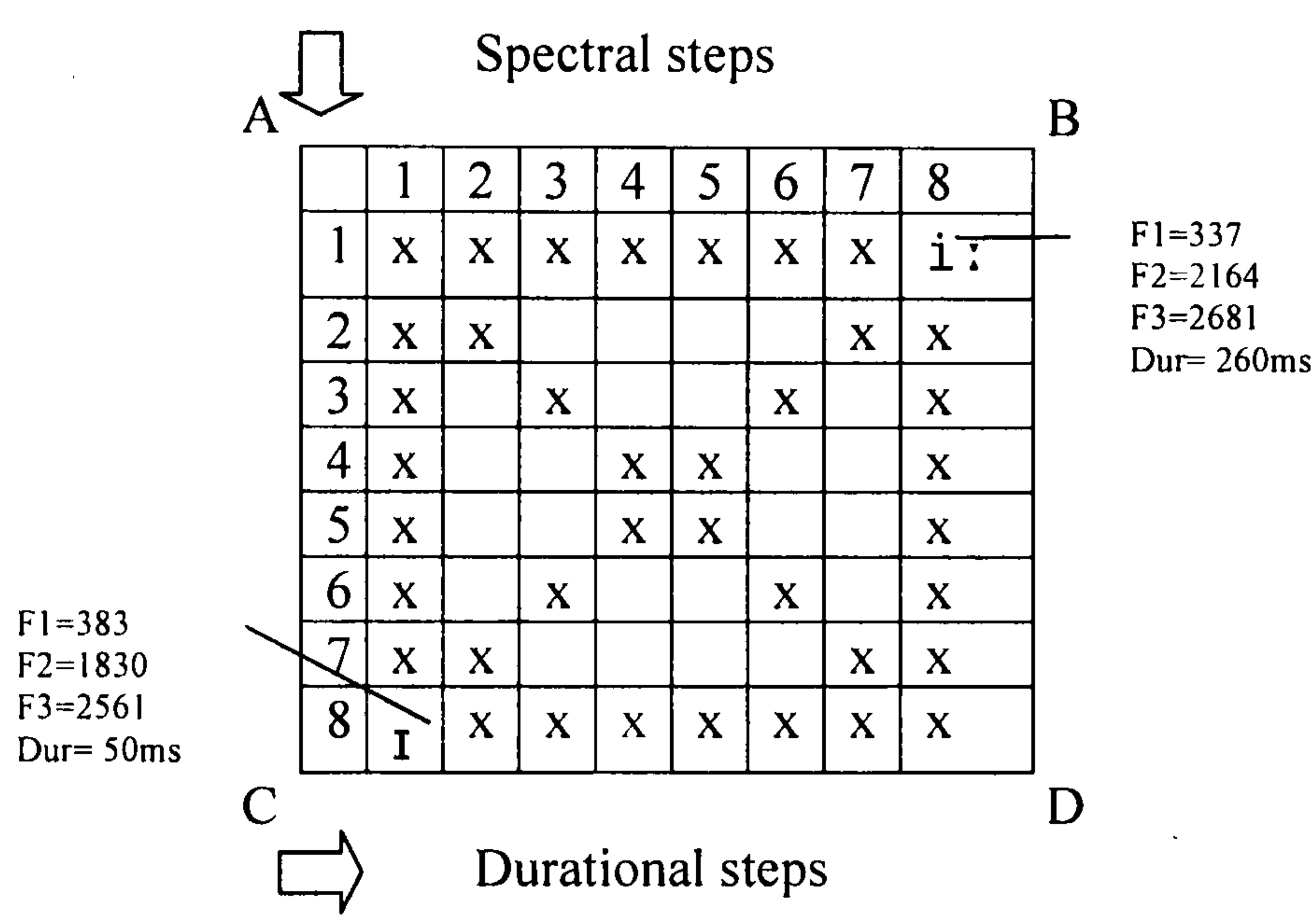
The stimuli generated for the experiment are summarised in Figure 5.2. As can be observed, not all possible combinations were explored, owing mainly to time constraints.

Fig.5. 1 Linear interpolation and parabolic morphing of a long-short continuum in order to adapt spectral values to differing durations



In the grid in Fig. 5.2, the shaded cells represent the end points /i:/ and /ɪ/. The upper right corner cell represents a long vowel /i:/ with maximal duration and formant frequency measures (the longest token produced naturally with high F2 and low F1) while the lower left corner cell represents a short vowel /ɪ/ (the shortest vowel produced with the most centralised F1 and F2 measurements). Then the formant and duration values for the other steps along the continua (marked with x) were computed using linear interpolation.

Fig.5. 2 /i:/ - /ɪ/ continua



Having generated the formant and pitch transitions, the next step was to use them with a speech synthesiser to create the audio data for the perception experiment. The synthesis method that lent itself to the material was the source-filter approach (see section 2.1 in Chapter Two). This method used the modelled F0 transitions to specify the pitch of the vocal source, whilst the formant transitions determined the characteristics of the filter. Praat was chosen to carry out the synthesis because of its easily implemented synthesis process and its scripting capabilities. Matlab (<u>www.mathworks.com</u> – version 5.3) was used to generate the pitch and formant transitions, and a separate data file was created for each of the 40 stimuli shown in Figure 5.2 above. Each file consisted of a matrix containing values for time, F0, F1, F2, F3, and formant bandwidths for F1 to F3 across each of the stimuli. The bandwidths are essential parameters in the synthesis process and were calculated using Fant's (1972) empirical formulae for B1, B2 and B3. A Praat script was written[20] which read the contents of each of the data files and then used this information to generate each of the vowels. In the synthesis process the 4th and 5th formants (F4 and F5) were specified with fixed values of 3570 Hz and 4500 Hz respectively, and their associated bandwidths with fixed values of 55 Hz and 100 Hz.

Rather than attempting to synthesize the /d/ and /b/ consonants, it was decided to use the concatenation synthesis technique for combining the synthesised vowels with real instances of the consonants. These were extracted from the recording used to determine the initial formant and pitch measurements. Each occurrence of /d/ and /b/ was tested with various synthesised vowels, and the most natural sounding instance of each was selected. The same instances of /d/ and /b/ were combined with each of the synthesised vowels. However, the hold phase preceding the /b/ was artificially generated and linearly scaled across the stimuli from 0.03 seconds to 0.13 seconds according to the duration of the vowel. This produced more natural sounding tokens than a fixed-duration hold phase. This further stage of generating the stimuli was done within the same Praat script as the main vowel synthesis.

A pilot perception test was carried out with the generated stimuli in which the final /b/ consonant was not present. This was done because initial concatenation attempts resulted in unnatural sounding stimuli when a final /b/ was added. In some cases the semblance of a /b/ was present simply from the formant transitions. However, after further experimentation with different instances of /b/ and differing hold phases, it

[20] Many thanks to Philip Harrison for writing this script.

became apparent that natural sounding stimuli could be achieved without a final /b/. This gave the utterances a more natural sound overall.

In order to add a further degree of realism to the synthesised stimuli, each token was subject to artificial reverberation within the sound editing software Soundforge (www.sonycreativesoftware.com – version 9). This was done by way of a batch process which automatically applied the reverberation to each file.

The final stage before the perception experiment consisted of randomising the order of the stimuli for presentation to the test subjects. Six of the stimuli were repeated within the random sequence to allow for some assessment of intra-subject variability. A Praat script was written to perform the randomisation process. In what follows, a detailed account of each set of continua is given.

5.1.2.1 Combined cues continua

This first set consists of two diagonal continua (i.e. CB and AD) each consisting of eight elements (see Figure 5.2). As mentioned above, C and B represent the endpoints /ɪ/ and /iː/ respectively. The diagonal continuum CB starts from point (r8 c1) (r refers to the row and c to the column) which has the values 383 Hz for F1, 1830 for F2, 2561 for F3, and 50ms for the duration and ends at point (r1c8) which has the following values: 337Hz, 2164Hz and 2681Hz for F1, F2 and F3 respectively. It also has the duration value of 260ms.

On the other hand, the diagonal continuum AD starts from stimulus (r1c1) which has the spectral values 337Hz, 2164Hz and 2681Hz for F1, F2 and F3 respectively and a duration value of 50 ms, and ends at stimulus (r8 c8) with the following spectral values: 383 Hz for F1, 1830 for F2, 2561 for F3 and the durational value of 260ms. However, in this continuum the durational values and formant values are reversed. That is, a short duration is given to the starting point of the continuum which has the formant values of a long vowel /iː/ (A), and a long duration is given to the end point of the continuum which has the formant values of the short vowel /ɪ/ (D). This set of continua was used to conduct an identification test in order to gather the perceptions of the vowel contrast and the category boundary of the two vowels.

5.1.2.2. Durational continua

The second set of continua are the durational continua where it was intended to test the participant's reliance on durational cues in perceiving vowel contrasts. Therefore, in each continuum of this set, spectral values were kept constant while the duration was manipulated in the eight stimuli of each continuum. These durational continua are AB and CD (see Figure 5.2 above).

The AB continuum has the following spectral values: 337 Hz, 2164 Hz and 2681 Hz for F1, F2 and F3 respectively. On the other hand, the values for the CD continuum are 383 Hz, 1830 Hz and 2561 Hz for F1, F2 and F3 respectively. The duration values for both continua ranged from 50ms to 260ms.

5.1.2.3. Spectral Continua

The third set that is generated are the spectral continua where it is intended to test the participant's reliance on spectral cues in perceiving vowel contrasts. Therefore, durational values are kept constant in each continuum and the quality of the stimuli is manipulated in the eight stimuli of each continuum. This set consists of two continua. They are AC and BD (see Figure 5.2 above).

The approximate natural values of the vowels produced by the native speaker in the mono-syllabic words /dɪb/ and /di:b/ were synthesised to make the various continua. The formant values for both continua ranged from 337 Hz, 2164 Hz and 2681 Hz to 383Hz, 1830Hz and 2561Hz for F1, F2 and F3 respectively. The durational value for BD was approximately the longest durational value of the vowel /i:/ (260ms), while that for AC (50ms) was the shortest that could feasibly be chosen while still enabling the researcher to manipulate its dynamic formant frequencies.

5.1.3. Participants

The participants were 20 native listeners of LA who came from the same town as those who participated in the production task. Table 5.2 shows their ages and foreign language experience.

Table 5. 2 Participants' age and English experience (perception experiment)

Participants	Age	Foreign language Experience in years		
		Preparatory	Secondary	University/collage
1	18	3	3	1
2	19	3	3	0
3	21	3	3	1
4	23	3	3	2
5	21	3	4	1
6	41	3	3	0
7	30	0	3	0
8	28	3	3	1
9	22	3	4	2
10	33	3	0	0
11	38	3	3	4
12	32	0	3	1
13	27	3	3	0
14	24	3	3	3
15	24	3	3	1
16	21	3	3	1
17	42	2	0	0
18	24	3	3	1
19	19	3	3	4
20	26	3	3	0

Optimally, these listeners should have been the same subjects who took part in the production task. However, because of the temporal gap between the two tasks it was not possible to find all those who participated in the production experiment, even though efforts were made to locate as many of them as possible for the perception task. None of the participants had histories of either hearing or speech problems.

Before performing the task, some demographic information, including place and date of birth, was gathered from the participants. This was to ensure that they had similar language backgrounds to those who participated in the production task. Next, participants were given an explanation of the study, and were presented with an introduction to the nature of their task. They were also asked to identify the words referred to by pictures in order to find out whether they recognised these words and could make a distinction between them. They were also presented with some vowel identification tasks in order to familiarise them with the real task. The perception task took place in a quiet room in the town of Rayaina where the participants lived.

5.1.4. Procedure

The participants performed three forced-choice picture identification tests. These consisted of listening to randomised tokens of [d_b] via a headphone and deciding whether the word they heard was /dɪb/ "*walk slowly*" or /diːb/"*wolf*" by clicking the correct picture on the screen using a computer programme called DMDX, which is designed to run stimulus/response experiments. As mentioned in section 5.1.2, the choice of these two words was based on mainly technical considerations. Therefore, these two words were selected for the experiment in spite of the fact that they do not belong to the same grammatical category. The use of pictures instead of written words was made to eliminate any effect resulting from the written forms on the participants' perception. The interval between each response and the presentation of the following stimulus was 60 seconds.

The first test was intended to determine the acoustic boundary between the two vowels from a perceptual point of view, and therefore both formants and duration values were manipulated in the set of continua used in this test. In the second continua of this set, durational values were coupled with a reversed spectral continuum leading to conflicting acoustic cues (e.g. short duration with spectral values of a long vowel /iː/ and long duration with spectral values of a short vowel /ɪ/). This was done to test which acoustic cue listeners would respond to if they had to make a choice.

In the second test durational values only were changed orthogonally from one stimulus to another. The aim was to investigate the participants' reliance on durational cues in perceiving these vowels. If the listener relied on durational cues, they would be expected to have a clear category boundary in the set of durational continua.

Finally, the third test consisted of presenting a set of two continua that aimed to examine the listeners' reliance on spectral cues in perceiving these vowels. Therefore, only spectral cues were manipulated in these continua while keeping durational cues consistent. If the subject relied on spectral cues, they would be expected to show a clear category boundary in the two spectral continua.

All three sets of the vowel continua were randomised, compiled in one file and presented to the participants using DMDX. After listening to each token, participants were asked to click on one of the two alternatives (/dɪb/ (*walk slowly*) and /diːb/ (*wolf*)); even if they felt uncertain, as this was a forced choice task. When the picture was clicked, the participant would hear the next word until all of the stimuli had been presented. The total number of stimuli was 40. However, these were divided into three files in order to give participants some rest after each file was played. The participants were also presented with four additional stimuli as examples at the beginning of each part in order to give them practice in how to use the programme. These initial four stimuli were not included in the analysis of the results. When the participants finished the task, they received a short message thanking them for their participation, followed by verbal thanks. DMDX writes a log file as it runs, containing input item numbers and responses so at the end of the task the results are ready for analysis.

5.1.5. Perception data analysis

The perception task went smoothly and was conducted easily. However, some participants admitted that some stimuli were tricky. The results of these tests for each set of continua (combined cues, durational and spectral continua) were dealt with and analysed separately. So the first task after gathering the data was to separate the results of all three sets of continua from each other by referring to the numbers of the stimuli in the matrix. As with the production experiment, the Microsoft Excel program was used to analyse the data and extract relevant statistics such as totals, means, percentages, standard deviations and significance levels.

5.2. Perception experiment results

This section focuses on three major points. First, an analysis and discussion of the results of the set of the diagonal continua is conducted. As mentioned before (see section 5.1.2.1), this set of continua was mainly designed to investigate the acoustic boundary between the two vowels /iː/ and /ɪ/ as they are perceived by Libyan native speakers.

Second, the set of spectral continua are analysed and discussed. These continua were designed in order to explore to what extent listeners rely on spectral cues in making a distinction between the two vowels /iː/ and /ɪ/. Finally, the results of the test involving the third type of continua, i.e., the durational continua, are dealt with. This type was generated in order to explore the effect of duration on the perception of these two vowels.

5.2.1. Combined cues continua analysis and discussion of results

The set of the combined cues continua consists of two continua. In these continua, both formant frequencies and duration were manipulated as shown in Figure 5.3.

Fig.5. 3 Diagonal /ɪ/ - /iː/ Continua (Shaded cells represent the stimuli of the continua which have decreasing/increasing durational and spectral values depending on the direction of each continuum)

⇨ Durational steps

	1 50ms	2 80ms	3 110ms	4 140ms	5 170ms	6 200ms	7 230ms	8 260ms
1	/iː/							/iː/
2								
3								
4								
5								
6								
7								
8	/ɪ/							/ɪ/

⇧ Spectral steps

However, in one of these continua the duration is reversed, giving long duration to /ɪ/ spectral values and short duration to /iː/ spectral values. Both the durational and spectral values used here are those values of the natural vowels that were synthesised to obtain these continua. The spectral values used at the endpoints of the continua were similar to the average values obtained from the production experiment for LA vowels /iː/ and /ɪ/[21].

However, while the minimal durational value used was similar to the average value obtained from the production task for the short vowel /ɪ/, the maximal value was considerably higher than the average value obtained from the production task for the long vowel /iː/. It was, in fact, similar to the maximal value obtained from the production task for the vowel /iː/. This is so since it was not possible to create eight durational steps which showed noticeable perceptual differences among them within the limits of the average durational values for both vowels obtained from the production task[22].

The first continuum of this set starts from the stimulus that has formant values of 383 Hz and 1830 Hz for F1 and F2 respectively and a duration of 50ms.This continuum ends at the stimulus that has formant values of 337Hz and 2164Hz for F1 and F2 respectively and the duration of 260ms as shown in Figure 5.4.

Fig.5. 4 Diagonal /ɪ/-/iː/ continuum (Shaded cells represent the stimuli of the continuum).

	1 50ms	2 80ms	3 110ms	4 140ms	5 170ms	6 200ms	7 230ms	8 (260ms
1								/iː/
2								
3								
4								
5								
6								
7								
8	/ɪ/							

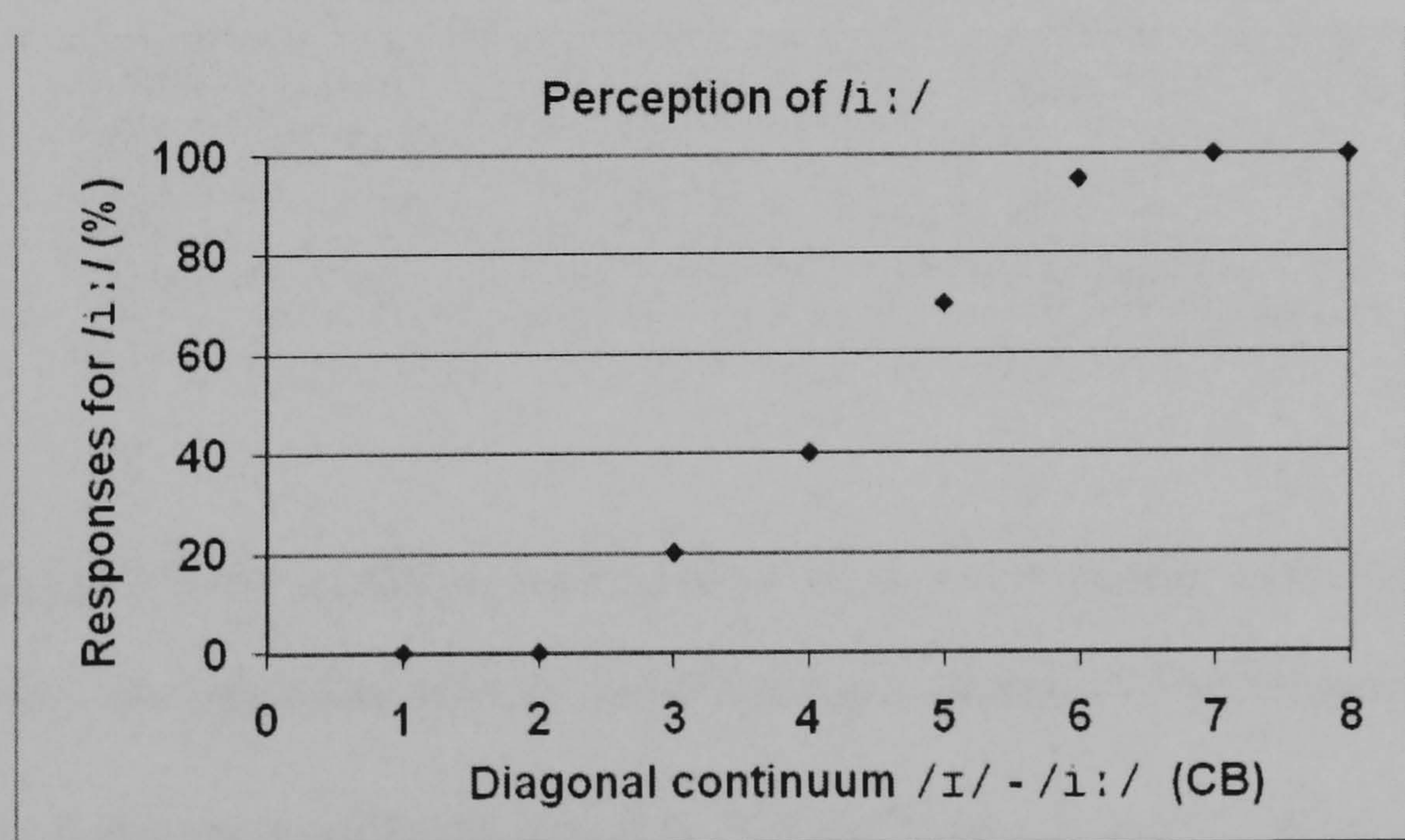
[21] The average formant values obtained from the production task were 342 Hz and 2214 Hz for F1 and F2 respectively for vowel /iː/, and 404 Hz and 1856 Hz for F1 and F2 respectively for the vowel /ɪ/.

[22] The average durational values obtained from the production task were 54ms and 138ms for vowel /ɪ/ and /iː/ respectively. The maximum durational value obtained for the vowel /iː/ was 245 ms and the minimum durational value for the vowel /ɪ/ was 22ms.

The results showing how the stimuli of this diagonal continuum were perceived by the 20 native participants are presented in Figure 5.5 below, indicating that the crossover point from one vowel to another varies from one speaker to another. However, the major shift from one vowel to another is between stimuli 4 and 5. While stimulus 4 was perceived as /ɪ/ by the majority of the participants (i.e. 60%), stimulus 5 was perceived as /iː/ by the majority (i.e. 70%). The variation between listeners in their crossover point supports the continuous nature of vowel perception frequently reported by researchers (e.g. Fry et al 1962). There is no clear-cut boundary between the two vowels when the perceptions of listeners are considered together, as shown in Figure 5.5 below.

These results are also consistent with "the perceptual magnet" proposed by Kuhl which refers to a certain location in the vowel space around which prototypes or the best instances (Rosch 1975) of the vowel as perceived by listeners gather. For example, all participants perceived stimuli 2 as /ɪ/ and not as /iː/ because it is nearer to the endpoint

Fig.5. 5 Perception of the stimuli of the combined cues continuum



that represents the perceptual magnet of the vowel /ɪ/. The same thing applies to the other side of the continuum. Stimulus 7 was perceived by all the participants as /i:/ and not /ɪ/ because it is nearer to the endpoint that represents the perceptual magnet of the vowel /i:/. However, when the distance between instances and perceptual magnets becomes bigger, their effect becomes weaker. Therefore, there is no 100% agreement among speakers as to what vowel stimuli 4 or 5, for example, belong.

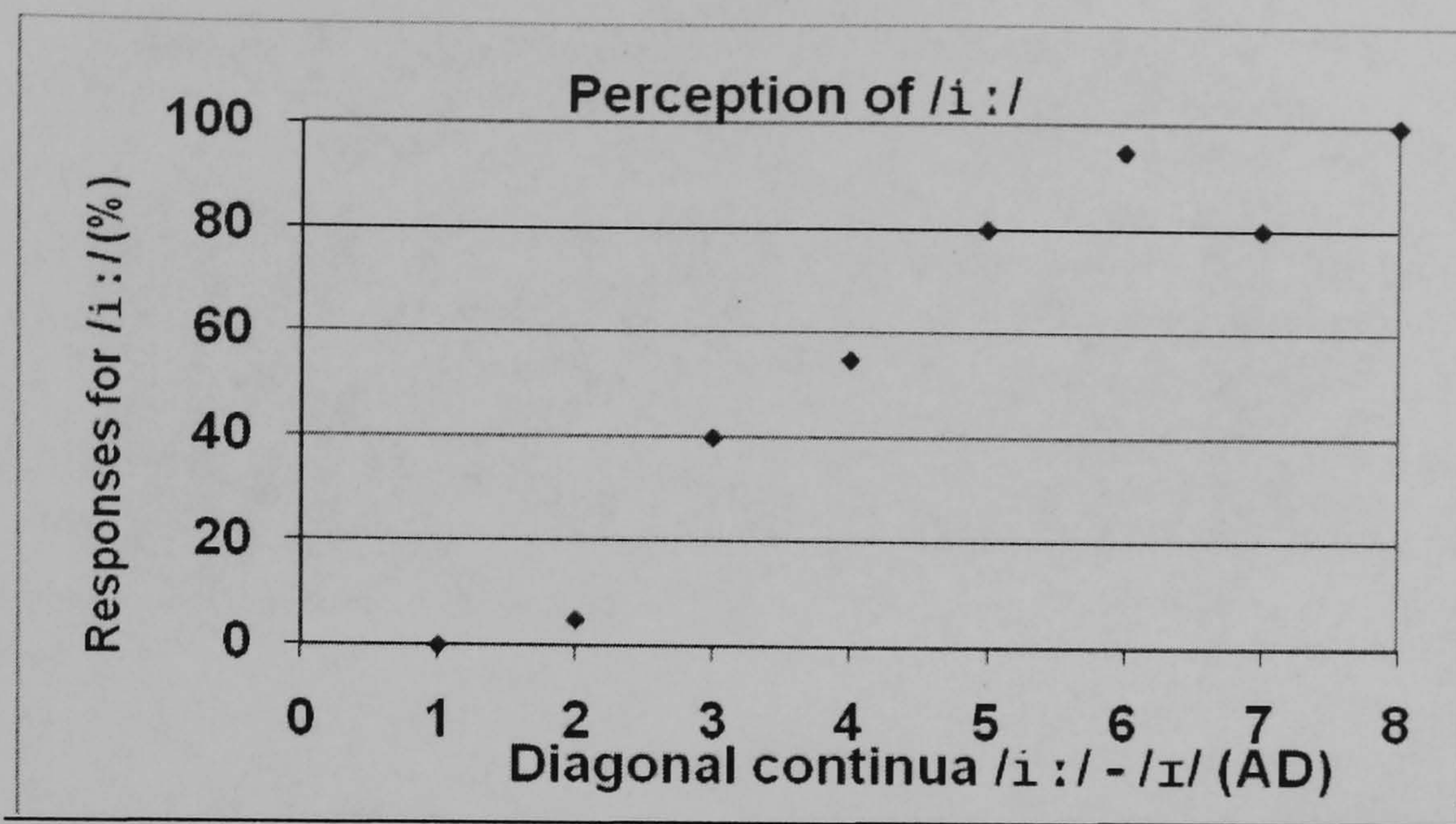
The second continuum of this set starts from the stimulus that has formant values of the vowel /i:/ and ends with the stimulus that has the formant values of /ɪ/. However, the duration of this continuum is reversed giving short duration to the stimulus that has /i:/ spectral values and long duration to the stimulus that has /ɪ/, with gradual durational steps between them as shown in Figure 5.6.

Fig.5. 6 Combined cues continuum /i:/-/ɪ/ with reversed duration (Shaded cells represent the stimuli of the continuum).

	1 50ms	2 80ms	3 110ms	4 140ms	5 170ms	6 200ms	7 230ms	8 260ms
1	/i:/							
2								
3								
4								
5								
6								
7								
8								/ɪ/

The results of how this continuum was perceived by the twenty speakers is shown in Figure 5.7.The Figure shows that the number of stimuli categorically perceived as /i:/ or /ɪ/ by all speakers is smaller than that in the first combined cues continua which was

Fig.5. 7 Perception of the second combined cues continuum /iː/-/ɪ/ with reversed duration



previously discussed. In this continuum only 1 stimulus was perceived as /ɪ/ by all listeners compared to 2 stimuli in the first continua. By contrast, only 1 stimulus was perceived as /iː/ by all listeners compared to 2 stimuli in the first continuum. As a result the shaded area which represents stimuli perceived as /iː/ or /ɪ/ by some speakers is bigger than that of the normal continuum as seen in Tables 5.3 and 5.4 below.

Table 5. 3 Perception of the first /ɪ/-/iː/ combined cues continuum

Stimuli	Perceived as /iː/		Perceived as /ɪ/	
	Participants	Percentage	Participants	Percentage
1	0	0	20	100
2	0	0	20	100
3	4	20	16	80
4	8	40	12	60
5	14	70	6	30
6	19	95	1	5
7	20	100	0	0
8	20	100	0	0

Table 5. 4 Perception of the second /ɪ/-/iː/ combined cues continuum

Stimuli	Perceived as /iː/		Perceived as /ɪ/	
	Participants	Percentage	Participants	Percentage
1	0	0	20	100
2	1	5	19	95
3	8	40	12	60
4	11	55	9	45
5	16	80	4	20
6	19	95	1	5
7	16	80	4	20
8	20	100	0	0

These results might suggest some role played by spectral cues in the perception of these vowels. When listeners were presented with unusual mixtures of spectral and durational cues, their perception was affected. That is more evident in the perception of stimuli number 7 which was perceived as /ɪ/ by 20% of speakers although it is contiguous to the endpoint stimuli representing the vowel /iː/. However, the general trend was maintained by all listeners. It seems that the major role is played by durational cues in distinguishing between the two vowels. How much listeners rely on any of these cues and how much this reliance weigh in their perception of these two vowels will be more obvious when the results of the durational and spectral continua are analysed and discussed.

Regarding the boundary between the two vowels, the crossover point at which perception shifts from one vowel to another by the majority of participants is between stimuli 3 and 4. The former was perceived as /ɪ/ by 60% of the participants and the latter as /iː/ by 55% of them. While this crossover point moves further to /ɪ/ as compared to the example of the first combined cues continua which was between stimuli 4 and 5, the shaded area is not divided equally by this cross-point. There are 4 stimuli beyond this point in the direction of /iː/ that were perceived as /ɪ/ by some listeners compared to only 2 stimuli in the opposite direction that were perceived as /iː/ by some listeners.

The differences between the perception of the first and second combined cues continua hint at some role that spectral values play in the perception of these two vowels. When the short vowel /ɪ/ was presented to listeners with the spectral cues of /iː/, there was no 100% agreement that a stimulus was /iː/ except for the 8th stimulus with values of 383 Hz for F1 and 1830 Hz for F2 (i.e. the spectral values of vowel /ɪ/). Similarly, when listeners were presented with the spectral values for /ɪ/ coupled with the /iː/ duration, there was no 100% agreement that a stimulus was /ɪ/ except for the first stimulus which had values of 337 Hz for F1 and 2164 Hz for F2 (i.e. /iː/ spectral values). However, this role might not be that important, especially when more important cues are available for listeners to rely on in their distinction between these two vowels; that is, the durational cues. As stated before, the role of spectral and durational cues in the perception of the two vowels /iː/ and /ɪ/ will become clearer when the results of the durational and spectral continua are analysed and discussed in the following two sections.

5.2.2. Spectral continua results: analysis and discussion

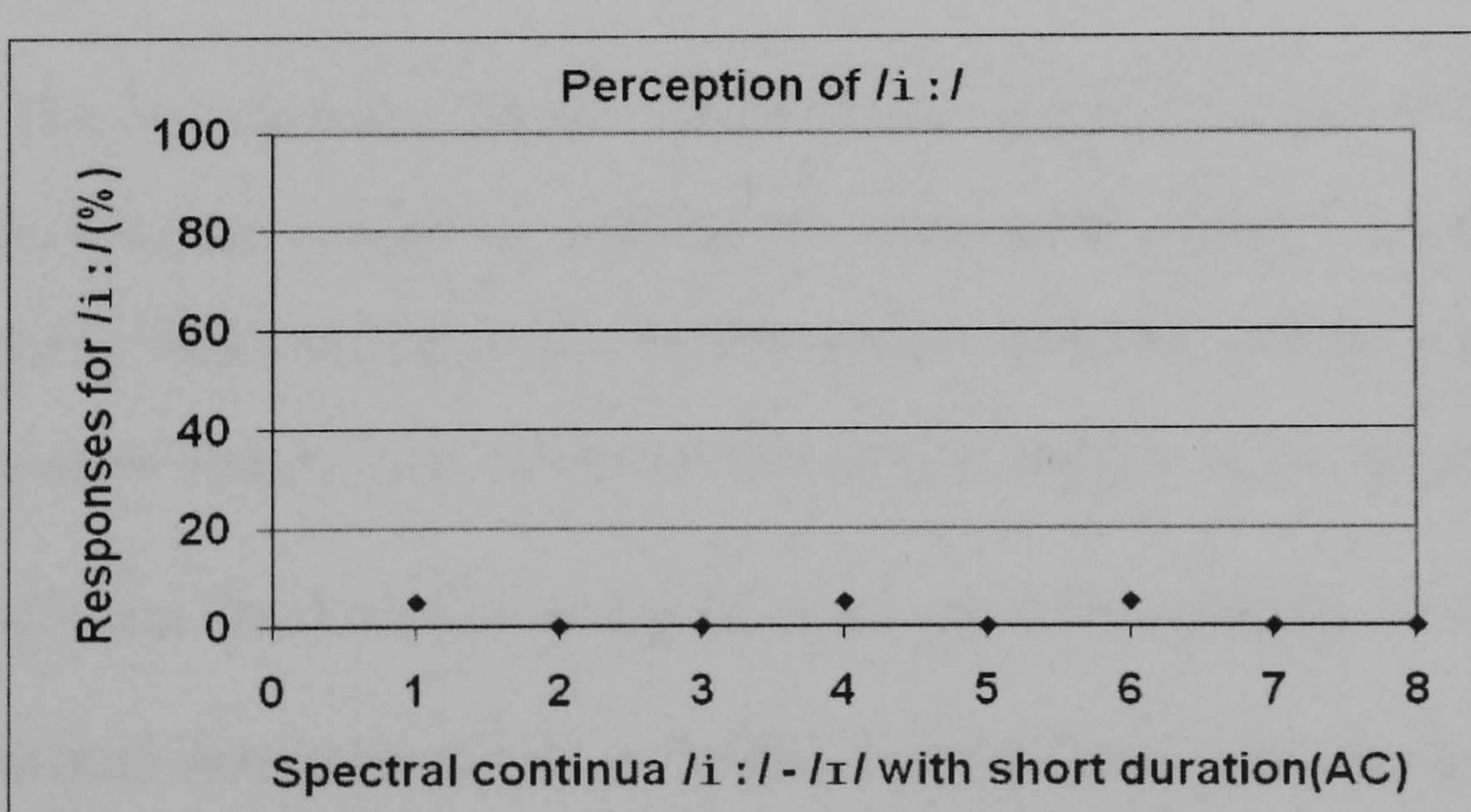
The set of spectral continua consists of two continua. In both continua, formant frequencies are manipulated while duration is kept the same for all stimuli of each continuum as shown in Figure 5.8 below.

Fig.5. 8 Spectral continua

	1st continuum 50ms	2nd continuum 260ms
1	/iː/	/iː/
2		
3		
4		
5		
6		
7		
8	/ɪ/	/ɪ/

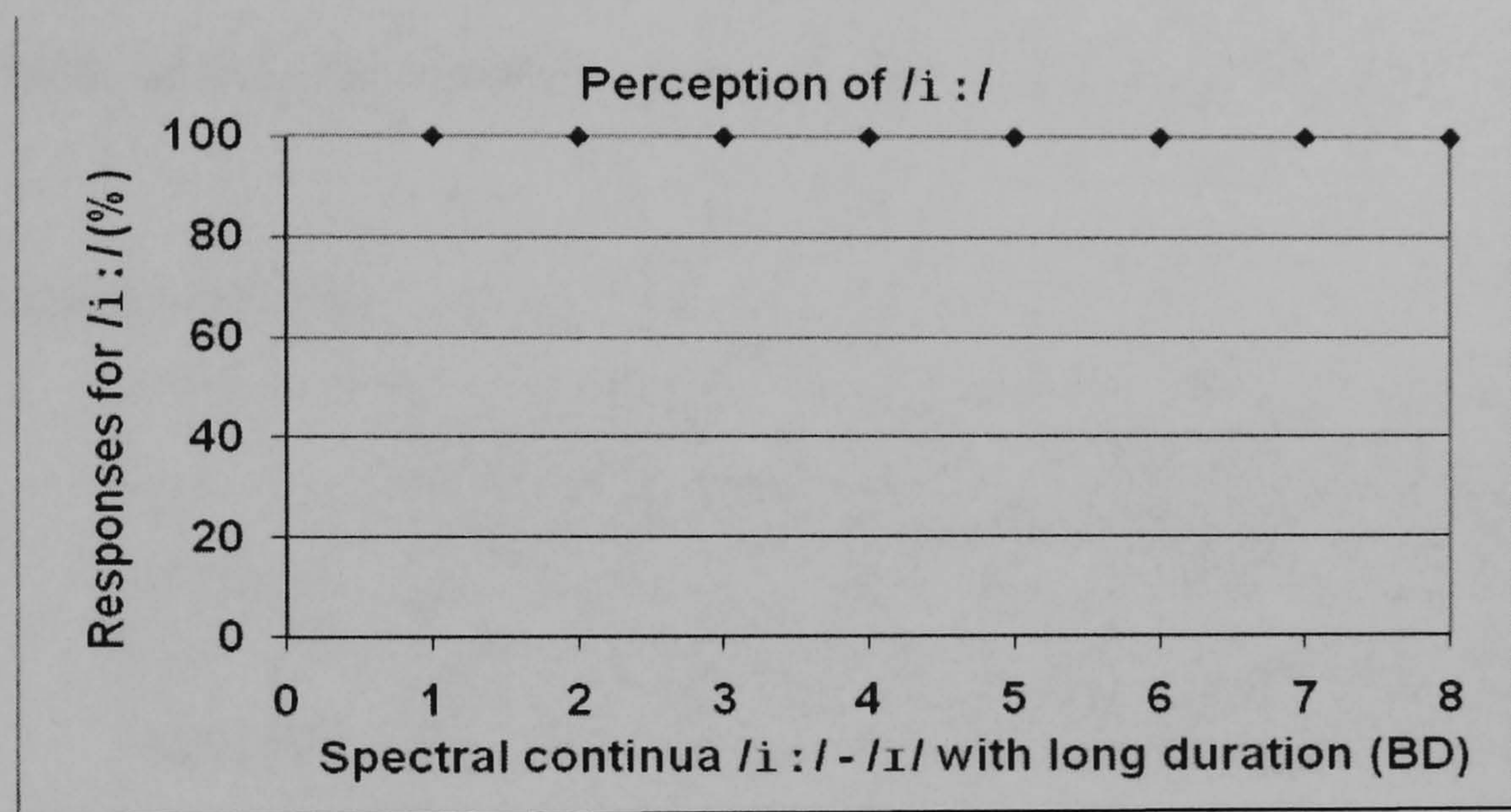
The perception results for the stimuli in the spectral continuum, which has a short duration, are shown in Figure 5.9 below. The results show that duration has greater influence than spectral cues in the perception of the stimuli of this continuum. However, some stimuli were distinguished according to their quality rather than their duration by some speakers. For example, stimulus 1 was perceived as /iː/ by one listener in spite of the fact that it has a short duration. Other listeners also perceived some stimuli as /iː/.

Fig.5. 9 Perception of the spectral continuum /iː/-/ɪ/ with short duration.



These results are supported by those of the second spectral continuum, which has a long duration, as shown in Figure 5.10, which confirms yet again that duration plays a major role in distinguishing between these two vowels. All stimuli of the spectral continuum which has a long duration were perceived as /i:/ by all listeners. These results show that spectral values are not important in perception when making a distinction between the two vowels /i:/ and /ɪ/.

Fig.5. 10 Perception of the spectral continuum /i:/-/ɪ/ with long duration



This is in harmony with the "desensitization" hypothesis (Bohn 1995) which maintains that when listeners fail to distinguish vowel contrasts based on spectral cues, they resort to durational ones because the latter are more prominent and easier to access. Bohn used the English /i-ɪ/ continuum with Spanish speakers whose language only contains /i/ and, therefore, duration does not play a role in their language. Nevertheless, these speakers were found to rely on duration in differentiating between the English vowels /i/ and /ɪ/ rather than spectral cues because duration cues varied in only three steps in the continuum whereas spectral cues varied in eight steps, which makes duration more prominent and easier to access than spectral cues.

5.2.3. Durational continua results: analysis and discussion

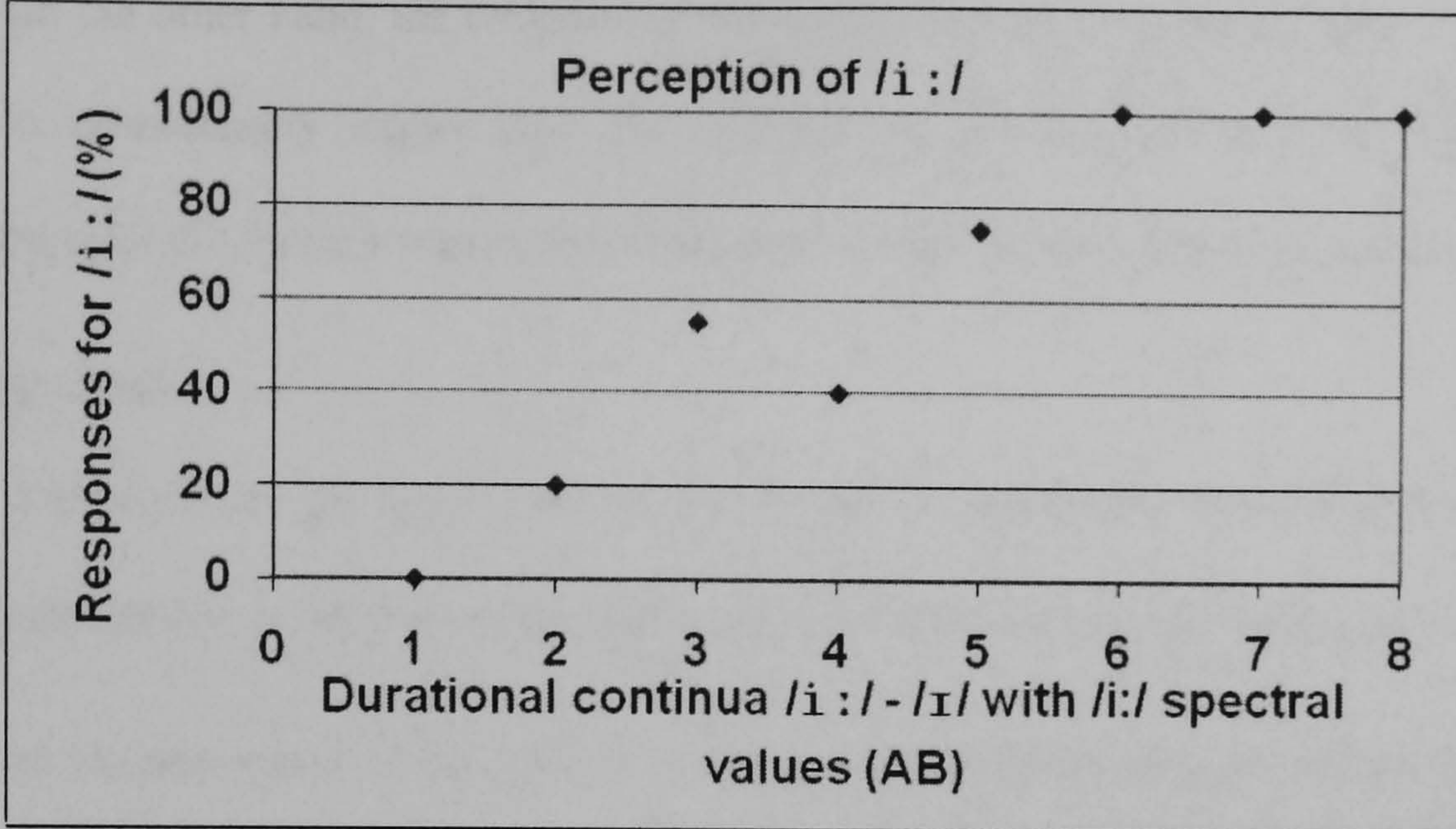
The set of the durational continua consists of two continua. In these two continua duration is manipulated while formant values are kept constant for all stimuli. However, the constant spectral values of one continuum are different from those of the other. While one continuum consists of the basic spectral values for the vowel /iː/, the other consists of the basic spectral values for the vowel /ɪ/, as shown in Figure 5.11 below.

Fig.5. 11 Durational continua

	1 50ms	2 80ms	3 110ms	4 140ms	5 170ms	6 200ms	7 230ms	8 260ms
1st continuum	/iː/	/iː/	/iː/	/iː/	/iː/	/iː/	/iː/	/iː/
2nd continuum	/ɪ/	/ɪ/	/ɪ/	/ɪ/	/ɪ/	/ɪ/	/ɪ/	/ɪ/

The results showing how the stimuli of the first durational continuum, which has the spectral values for /iː/, were perceived by the 20 native participants are presented in Figure 5.12 below.

Fig.5. 12 Perception of the durational continuum /i:/-/ɪ/ with /i:/ spectral values

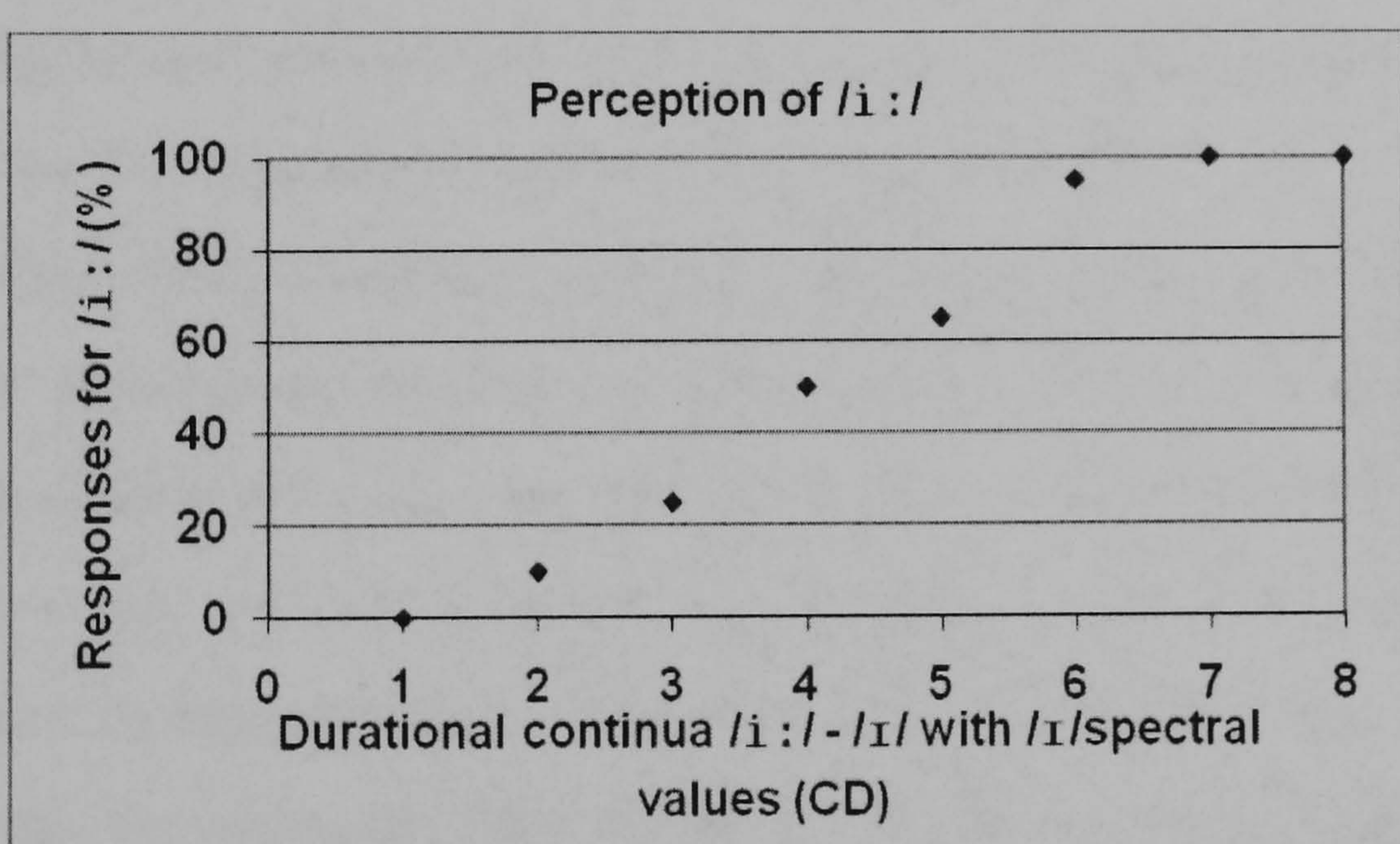


These results confirm that duration has major importance compared to spectral cues in the perception of the stimuli of this continuum. Although all stimuli have the spectral value of /i :/, listeners perceived short ones as /ɪ/ and long ones as /i :/. However, there was some confusion at the middle of the continua, specifically in stimuli 3 and 4. The number of listeners who perceived the vowel /i :/ rose from 20% in stimuli 2 to 55% in stimuli 3. However, instead of continuing to rise, it suddenly fell to only 40% per cent in stimuli 4 before it rose again to 75% per cent in stimuli 5. This fluctuation makes the vowel boundary difficult to decide, though it is certainly between stimuli number 2 and 5.

It is also apparent that the number of stimuli perceived as /i :/ by all speakers is greater than that perceived as /ɪ/. Three stimuli were perceived as /i :/ by all speakers compared to only one stimuli perceived as /ɪ/ by all speakers. This can be attributed to the fact that the starting point of the continuum has a duration value of 50ms, which does not differ much from the average value obtained from the production task for /ɪ/ which was 54ms. On the other hand, the endpoint of the continuum has a duration value of 260ms, which is considerably higher than the average value obtained from the production experiment for /iː/ which was 138ms. This extra length is often perceived by all listeners as a long vowel.

The results of the first durational continuum are supported by those of the second continuum for the durational set of continua which has the spectral values of /iː/. The results of the perception of the stimuli in this durational continuum are shown in Figure 5.13 below.

Fig.5. 13 Perception of the spectral continuum with /ɪ/ spectral values.



The results of this continuum confirm the results of the previous durational continuum, indicating that duration has major importance compared to spectral cues in the perception of the stimuli of the durational continuum. Although all stimuli have the spectral value of /ɪ/, listeners perceived short ones as /ɪ/ and long ones as /i:/. The vowel category boundary, which was difficult to decide in the previous durational continua, is now between stimuli 4 and 5 in this durational continuum. The midpoint between the two vowels here is within stimuli 4 which is divided equally between the two vowels; 50 % of the listeners perceived it as /i:/ while the other 50 % perceived it as /ɪ/. As with the previous durational continuum, the number of stimuli perceived as /i:/ by all speakers is greater than that perceived as /ɪ/ (by a ratio of 2:1).

5.2.4. Cue reliance and weighting

Having analysed all continua separately, it is time now to consider the overall effects of both durational and spectral cues. In order to do this, both durational and spectral reliance and weighting were calculated following Flege's (1997) procedure in calculating cue effects. The same procedure was followed by Escudero (2000, 2002).

Flege (1997) conducted a study on the effect of experience on non-native speakers' production and perception of English vowels. In his study, he compared an experienced group with a non-experienced group. In the perception experiment and in order to explore the effect of spectral cues and duration cues on the participants' perceptions, he calculated "spectral effect scores" and "durational effect scores" for the two groups. For example, the spectral effect scores were obtained by subtracting the percentage of responses given to the /ɪ/ endpoint from the percentage of responses given to the /i:/ endpoint. Then, the resulting scores were added together and averaged over the total number of stimuli in the durational continuum. The temporal effect was calculated in the same manner, by subtracting the percentage of responses given to the stimuli for the long vowel from those given to the short vowel, averaging the sum over the total number of stimuli in the spectral continuum. This procedure was also followed by Escudero (2000) who investigated the acoustic cue weighting in the perception of Scottish tense/lax vowels by Spanish speakers and compared it to that of native speakers of English.

The same procedure was used in this study to measure cue reliance in the perception of /ɪ/ and /iː/ vowels by Libyan speakers. The durational reliance score was calculated by subtracting the percentages of /iː/ responses given to stimuli of the short durational continuum from those given to stimuli from the long durational continuum, averaging the total number resulting from these subtractions over the total number of stimuli which is 8. Similarly, the spectral reliance score was calculated by subtracting the percentages of /iː/ responses given to stimuli of the continuum which has /ɪ/ spectral values from those of the continuum having /iː/ spectral values, before averaging the total number resulting from the subtractions over the total number of stimuli which is 8.[23]

Fig.5. 14 Durational and spectral /ɪ/ - /iː/ Continua

Durational steps

A B

	1	2	3	4	5	6	7	8
1	x	x	x	x	x	x	x	x
2	x							x
3	x							x
4	x							x
5	x							x
6	x							x
7	x							x
8	x	x	x	x	x	x	x	x

C Spectral steps D

[23] For example the durational reliance was computed using the following operation: (((BD1-AC1)+((BD2-AC2)+((BD3-AC3)+((BD4-AC4)+((BD5-AC5)+((BD6-AC6)+((BD8-AC8))/8 = durational reliance (See figure 5.14).

Cue weighting, on the other hand, was calculated by dividing each reliance score by the sum of the reliance scores for both spectral and durational cues. Again, this procedure was also followed by Flege (1997) and Escudero (2000). Table 5.5 shows both cue reliance and cue weighting results obtained in this study from the perception experiment.

Table 5. 5 Cue reliance and cue weighting results

	Durational scores	Spectral scores
Reliance	96.25%	05.62%
Weight	94.48%	05.52%

It is obvious from Table 5.5 that duration has significantly higher reliance and weighting scores than spectral values. In fact, spectral values amount to only 5.62% of the reliance score and even less than that for the weighting score.

5.2.5. Inter-listener and intra-listener variation.

Intra-listener variation was found in five listeners from the 20 who participated in the study. These listeners were numbers 5, 10, 12, 13 and 18, as shown in Figure 5.15 below. What is meant by variation here is that a participant's recognition of a certain vowel fluctuates from one stimulus to another. For example, listener 5 recognised stimuli numbers 1 and 2 as /ɪ/ and then stimuli numbers 3 and 4 as /iː/. His perception shifted to /ɪ/ in stimulus 5 before switching to /iː/ in the remaining stimuli. Similarly, listener 18 recognised the first two stimuli as /ɪ/ and then he perceived stimulus 3 as /iː/. His perception shifted to /ɪ/ in stimulus 4 before finally switching to /iː/ in the rest of the stimuli. These variations within speakers might have been caused by confusion or lack of attention during the test.

Variations among speakers are also obvious. For example, while the vowel boundary for listeners 3 and 16 was between stimuli 2 and 3, it was perceived between stimuli 4 and 5 by 8 other listeners as, for example, listeners 1, 7 and 20. However, the crossover point was made between stimuli 5 and 6 by some other listeners, including numbers 4, 9 and 17.

Assuming that the vowel boundary is between stimuli 4 and 5, it can be observed that 7 /iː/ stimuli were perceived as /ɪ/ whereas, on the other side of the boundary, 11 /ɪ/ stimuli were perceived as /iː/. However, most of these reversed responses are in the middle of the continua where neither durational nor spectral cues were sufficient to define the vowel identity with certainty.

Fig.5. 15 Inter-speaker variation in the perception of the first diagonal continuum (x axis= stimuli, y axis= listeners)

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ
2	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ
3	ɪ	ɪ	iː	ɪ	iː	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	ɪ	iː	ɪ	iː	ɪ	ɪ
4	ɪ	iː	iː	ɪ	iː	ɪ	ɪ	ɪ	ɪ	ɪ	iː	iː	iː	ɪ	ɪ	iː	ɪ	ɪ	iː	ɪ
5	iː	iː	iː	ɪ	ɪ	iː	iː	iː	ɪ	iː	iː	ɪ	ɪ	iː	iː	iː	ɪ	iː	iː	iː
6	iː	iː	iː	iː	iː	iː	iː	iː	iː	ɪ	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː
7	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː
8	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː	iː

These results support the continuous nature of vowel perception along a continuum. Listeners continue to realise that stimuli belong to one vowel or the other despite the gradual change in the formant and durational values which are different from those of the vowels at the endpoints. However, participants vary in terms of the decisive boundary between the two vowels, with no agreement among speakers as to where the boundary is between the two vowels.

Variability in perception has been reported by some researchers. In a study by Hombert & Puech (1984) on Swahili, it was found that listeners have different perceptual areas and, therefore, the same vowel stimulus was assigned to the same or a different vowel category by different listeners. A possible source of variation might be the differences among listeners in their experiences, expectations and language backgrounds.

5.3. Summary

The perception of LA /ɪ/ and /iː/ was investigated in this chapter. The aim was to find out whether or not the cues found significant in production are also important in perception. The results show that durational cues are more important than spectral cues when these vowels are perceived. In fact spectral cues have little significance and participants mainly relied on duration for making a distinction between /ɪ/ and /iː/. Moreover most of the participants were able to draw a boundary between this pair of vowels although this boundary varied from one listener to another. In the following chapter, the research questions are reviewed in light of these results. The limitations of the study and recommendations for future work are also given.

CHAPTER VI

GENERAL DISCUSSION AND CONCLUSIONS

6.0. Introduction

In this chapter, the research questions of the study are reviewed and evaluated in the light of the results obtained. This is accomplished by presenting each question, discussing the results relating to that question and comparing them to findings for other Arabic dialects. Links to theoretical assumptions and universal linguistic issues are also made. But first, a brief reminder of the aims and the procedure of the present research is given.

6.1. Aims and procedure of the research in brief

This study aimed at exploring the vocalic system of LA and providing a comprehensive description of the vowels contained in the system. The motivation was the lack of empirical work on LA. It was hoped that the study would constitute the first experimental investigation of the phonetic and phonological difference among the vowels contained in the vocalic system of the dialect and establish a basis for further investigation. Another goal was to find out to what extent the findings of this study are consistent with findings from other Arabic dialects. A comparison of this type is necessary to explore the characteristics these dialects share.

To accomplish these aims two studies were carried out. One related to vowel production and the other was concerned with vowel perception by native speakers. The production task involved the use of 20 native speakers producing the vowels available in the dialect in monosyllabic words embedded in carrier sentences. Each speaker produced

160 tokens and the total number of tokens for the 20 speakers amounted to 3200 words. F1 and F2 frequencies and vowel durations were measured. Efforts were made to ensure the reliability of measurements by double checking them.

Based on the results of the production task, a perception experiment was also conducted. Only two vowels, namely /iː, ɪ/, were investigated in the perception experiment due to time and space limitations. This experiment aimed to find out what cues are most important when this long/short contrast is perceived. Another aim was to investigate whether or not these two vowels are perceptually separate in spite of their qualitative overlap in the vowel space. Therefore, three sets of vowel continua, each containing eight stimuli, were designed leading to a total of 40 stimuli for the three sets. As in the production task, a sample of 20 native listeners was involved in the perception experiment. The perception task involved 20 native listeners who performed a forced-choice picture identification task during which they heard randomised /d_b/ tokens in which duration and/or spectral values were manipulated and were asked to decide whether the stimulus they heard was the lexical item /dɪb/ or /diːb/. Again, the results were double-checked to ensure reliability.

6.2. Review of the research questions

The research questions, first presented in Chapter One, section 1.4, consist of one major question followed by four specific sub-questions as follows:

Main research question

What is the vowel inventory of Libyan Arabic and how are the vowels contained in this inventory produced and perceived by native speakers of the dialect?

Specific questions

1. What is the phonological inventory of vowels available in LA?
2. How are these vowels realised qualitatively and quantitatively in this dialect?
3. Which vowel acoustic cues do native speakers of LA respond to in perception?
4. To what extent does perception of these vowels relate to the acoustic cues that are found relevant in production?

Each of these specific questions is now answered before turning to the more general conclusions relating these sub-questions to the main question.

Q1. What is the phonological inventory of vowels available in LA?

The review of the relevant literature has shown that the Libyan vocalic system contains eight vowels. These eight vowels constitute the vocalic systems of all Libyan dialects studied so far including ZA (Abumdas 1985), DA (Aurayieth 1982), SA (Botagga 1991) and TA (Muftah 2001). However, the realisations and phonological distribution of these vowels may vary from one dialect to another (see section 3.2.1 in Chapter Three). Traditionally these vowels are symbolised as /i:, i, u:, u, e:, o:, a:, a/. It is apparent from these symbols that the qualitative difference between short and long vowels is disregarded. The failure to use experimental techniques in previous research prevented this difference from being recognised (more discussion of this difference between short and long vowels is given in the answer to the second research question, when the realisations of these vowels are discussed).

The same number of vowels is also found in other Arabic dialects: for example, Jordanian Arabic (Barkat-Defradas 2003) and Egyptian Arabic (Gairdner 1925, Cowan 1970, Norlin 1987). However, because most of the studies conducted on Arabic vowels have been concerned with comparing dialects to the standard variety, some vowels which are unique to the dialects were not included, such as /e:/ and /o:/. One other reason is that these two vowels do not have short counterparts to be compared with. According to some researchers these two dialectal vowels originated from the MSA diphthongs /ai/ and /au/ respectively. However, the availability of these two diphthongs alongside the two long vowels /e:, o:/ in some dialects, for example in LA, casts doubts on this view.

From a universal point of view, the LA inventory is considered one of the most common inventories in that it includes /i, e, a, o, u/, which constitute a model vowel inventory in the world's languages (Lang & Ohala, 1996).

Q2. How are these vowels realised qualitatively and quantitatively in this dialect?

The acoustic and auditory features of these vowels were discussed in detail in the previous two chapters. Quantitatively, LA vowels can be classified in two categories: long and short. The former includes /i:/, /u:/, /e:/, /o:/ and /æ:/ while the latter comprises /ɪ/, /ʊ/, and /ə/. The results of the production study have shown that long vowels are more than twice as long as short vowels. Table 6.1 shows their mean duration.

Table 6. 1 Duration means of LA vowels in milliseconds

/i:/	/ɪ/	/u:/	/ʊ/	/e:/	/o:/	/æ:/	/ə/
138	54	148	64	156	154	150	63

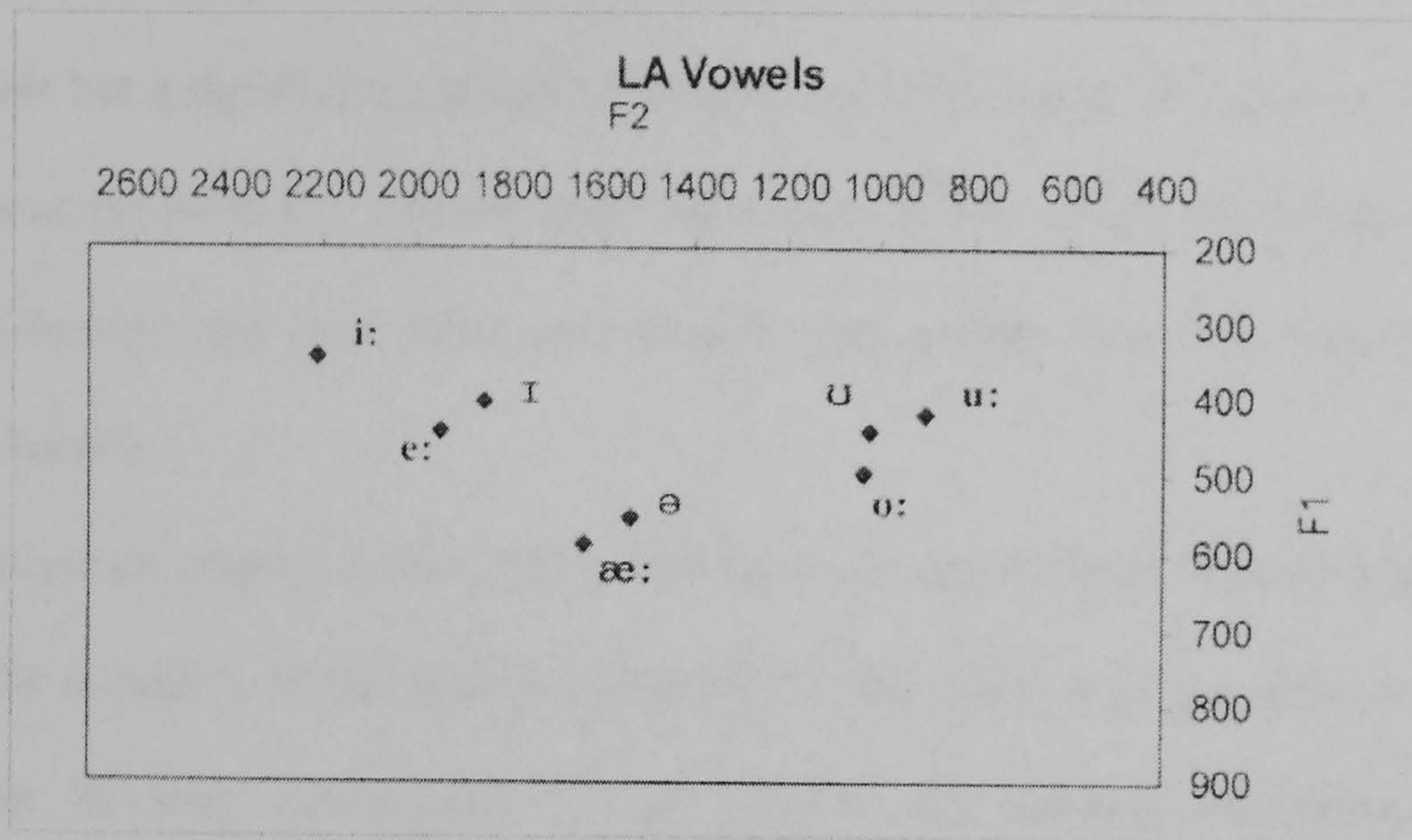
The means of the duration of long vowels were found to be significantly longer than those of short vowels. The mean for the short vowels together is 60 ms whereas that for long vowels is 149ms and the ratio between the two groups is 0.40. This durational ratio is the same as the one found by Algamdi (1998) for Egyptian Arabic and is similar to that of Sudanese (0.41) (Alghamdi 1998). It is, however, considerably different from those found for Saudi (Alghamdi), Iraqi (Al-Ani 1970), Jordanian (Mitleb 1984) and Gulf Arabic (Hussain 1985) which are 0.51, 0.50, 0.65 and 0.56 respectively.

The durational difference between high front vowels and their low and back counterparts was also found to be significant. Low vowels are significantly longer than high front vowels due to the extra time needed for lowering the jaw when low vowels are produced (Lindblom 1968; Lehiste 1970). High back vowels were unexpectedly longer, which can be attributed to contextual factors. Most of the consonants before and/or after these vowels are produced at the front of the mouth, which requires the tongue to move over a longer distance from/to the position of the vowel. This longer distance requires a longer duration.

Contrary to the traditional view that was held for a long time by some Arabic linguists, the group of short vowels also show significant qualitative difference from their long counterparts in LA and in other Arabic dialects. The results of the production experiment have shown that the F1/F2 plots for short vowels /ɪ, ʊ, ə/ and their long counterparts /iː, uː, æː/ seem to suggest that short vowels may be more centralised than their long counterparts, as can be seen from the representation of the two vocalic groups on the formant planes shown in Figure 6.1 below.

The effect this has on each vowel is the following: the short vocalic category represented by /ɪ/ seems to be more retracted and lower than its long counterpart /iː/. By contrast, the short vowel /ʊ/ is somewhat lower and more to the front than its long counterpart /uː/. Also, the short vowel /ə/ is higher and further back than its long

Fig.6. 1 Average means of LA vowels (Hertz)



counterpart /æː/. Figure 6.1 also shows that the qualitative difference between /iː/ and /ɪ/ is greater than that between /uː/ and /ʊ/ or /æː/ and /ə/. The auditory analysis (see cAppendix 5) conducted also confirmed this qualitative difference between short and long vowels. Table 6.2 shows the average means of their formants.

Table 6. 2 Average means of LA vowels

	/iː/	/ɪ/	/uː/	/ʊ/	/eː/	/oː/	/æː/	/ə/
F1	342	404	416	443	442	496	588	555
F2	2214	1856	907	1026	1950	1037	1641	1541

It should be emphasised, however, that these vowels exhibit considerable variation within vowel categories both within and between speakers, as the detailed analysis and discussion of the results in Chapter Four have shown.

This significant qualitative difference was also found between all short vowels and their long counterparts in other Arabic dialects. These include Tunisian (Belkaid

1984) and Saudi (Algamdi 1998). However other dialects exhibit this difference in some vowels only. For example, Al-Ani (1970) found comparable formant values for front and back vowels but a significant difference between short and long low vowels. The opposite situation was found in CA (Newman & Verhoeven 2002), where the difference between high long vowels and their short counterparts was greater than that between long and short low vowels.

Moreover, some Arabic dialects that have some difference between short and long low vowels exhibit a rather odd phenomenon. Some show that the short low vowel is lower than its long counterpart. These include, for example, Tunisian, Jordanian, Lebanese (Abou Haidar 1994), Egyptian (Alghamdi 1998) and Cairene (Newman 2002). Others show that the low long vowel is more retracted than its short counterpart as, for instance, in Iraqi (Al-Ani 1970), Syrian, Jordanian, Saudi, Qatari, Emirati (Abou Haidar 1994) and Sudanese (Algamdi 1998; Abou Haidar 1994).

This complicated picture regarding the qualitative status of Arabic vowels implies that it might not be possible to use the earlier suggested symbols for all Arabic dialects. Rather, each individual dialect should be treated separately and the symbols used should be consistent with the vowel qualities each dialectal system has. Based on this and having found a significant qualitative difference between long and short vowels in LA, the symbols /ɪ, ʊ, ə/ for short vowels as opposed to /iː, uː, æː/ for long vowels seem to be plausible. However, the production of speakers provides only half of the picture. This picture only becomes clearer when we consider the perception of these vowels and the role played by acoustic cues in this perception, which is dealt with in the answer to the following research question.

Q3. Which vowel acoustic cues do native speakers of LA respond to in perception?

Since considerable variability was found in the production of the vowels by all the speakers, there was some overlap between vowels in the vowel space. This overlap is attributed to factors such as differences among speakers in their vocal tract size, speech rate, idiosyncratic differences, and the context in which these vowels occurred in addition to some natural variability that might result in different realisations of these vowels between as well as within speakers.

The aim here was to investigate to what extent listeners rely on durational and spectral cues in recognising short and long vowels. Since it was not possible to investigate all overlapping vowels due to limited space and time, the perception experiment concentrated on data from the high front vowels /iː/ and /ɪ/.

The results of the perception experiments have shown that listeners are capable of drawing a boundary between the two overlapping vowels. However, there was variation among speakers regarding the position of the boundary along the /iː/-/ɪ/ continuum. These results support the continuous nature of vowel perception. There is no clear-cut boundary between the two vowels when the perceptions of all listeners are considered as a result of the variation among listeners in their choice of the boundary. However, Listeners continue to recognise the small changes in the F1 and F2 values of vowels as a gradient continuum in the vowel space until they reach the boundary between the two when they shift from one vowel to the other.

Regarding listeners' reliance on spectral and duration cues in making distinctions between these vowels, the perception results have shown that the major factor in distinguishing between short and long vowels was their duration. When participants were presented with spectral continua where only formant frequencies were manipulated while duration was kept constant, they generally recognised the continuum with short stimuli as being /ɪ/ and that with long stimuli as being /iː/. However, when they were presented with durational continua where only duration was manipulated while spectral cues were kept constant, they generally recognised short stimuli in both continua as belonging to /ɪ/ and long ones as belonging to /iː/ irrespective of the fact that one continuum had /ɪ/ spectral values and the other had /iː/ spectral values. When the cue reliance of the participants was calculated, it was found that durational cues amounted to more than 96% of the total cue reliance.

As stated before, this finding is consistent with the "desensitization" hypothesis proposed by Bohn (1995). According to this hypothesis listeners resort to durational cues when spectral cues are not sufficient to make a distinction between vowels because the former are thought to be more prominent and easier to access than the latter.

Cross-linguistically, listeners differ in the types of phonetic cues that they consider more important for the perception of vowel contrasts. While duration is found in this study to be the major cue in making influencing distinctions between vowel contrasts in LA, in other languages such as, for example, English, spectral cues have been found to be more important. For instance, Escudero (2000) found that Scottish speakers relied exclusively on spectral cues in distinguishing between the tense and lax vowels /i- ɪ/. Another study was conducted by Makashay (2003) also found evidence that frequency cues are primary while durational ones are secondary in the perception of vowels. However, Makashay's study also suggests that there are dialectal and individual differences in the way English listeners use duration and frequency in the identification of vowels (see also Giegerich1992).

Q4. To what extent does the perception of these vowels relate to the acoustic cues that are found relevant in production?

When production and perception results are compared, two main observations are made. The first relates to variation. Both perception and production were found to vary considerably between speakers. Generally, short vowels were found to show more variability in the vowel space than long vowels. This is possibly due to the context in which these vowels occur (Frieda 2000). Unlike long vowels, short vowels are often undershot and therefore their formant steady state is never reached (Strange 1989, p.2082). Different adjacent consonants result in different degrees of this effect (Lindblom 1963).

As far as their quality is concerned, high front vowels are less variable than the remaining vowels, and high back vowels are the most variable ones. High front vowels being less variable can be attributed to the fact that most of the neighbouring consonants in the words elicited are anterior. That is, they are produced in the same area where these vowels are articulated. On the other hand, back vowels are more variable because they do not entail precise control of the tongue height, resulting in more variability in the production of these vowels (Beckman et al 1995).

This variation has consequences in both production and perception. In production it results in overlaps between vowels in the vowel space when the production of all speakers is considered. The same thing is found in perception, resulting in different boundaries between vowels made by different speakers. However, this overlap is unlikely to cause misunderstanding because listeners have an internal system that contains prototypes representing the vowels in the language inventory. When they listen to speech, listeners map the sounds they hear to these internal prototypes in a process called normalisation in which listeners exploit different types of information which are both internal and external to the vowels perceived (Nusbaum & Morin 1992). Moreover, listeners use information that is both internal and external to the speech signal. This includes phonetic (articulatory, acoustic and auditory), lexical, grammatical, social-indexical, and other stored knowledge. The role of all these levels of representation is to facilitate the perception process (Lindblom, 1996, pp.1687-89).

The second observation from the comparison between speech production and perception is related to qualitative differences between short vowels and their long counterparts. While this difference was found to be significant in production it was not significant in perception where listeners mainly rely on the quantitative difference in making distinctions between these vowels. It should be admitted, however, that this finding in perception was derived from comparing high vowels only. Other vowels were excluded from the experiment due to time and space limitations.

The question that arises now is, since qualitative differences are not important in perception, why speakers maintain them in production. A plausible answer could be that language is evolving and changing. However, linguistic changes take place over a long period. A time might come when these observations in production will also occur in perception.

This reliance on duration in the dialect under investigation derives from language-specific characteristics. Duration has been found to be a major distinguishing factor between vowels at least in the Arabic dialects that have been investigated so far. However different languages have different modes of implementation of these cues. English, for instance, exhibits more importance of spectral cues over durational cues (Escudero, 2000). It should be emphasised, however, that apart from this study scant research has been conducted on speech perception in Arabic dialects, and therefore any generalisation of the results should be made with caution. Furthermore, this study is only limited to one

Libyan dialect and, thus, further investigation of other Libyan dialects is necessary to ensure that these findings are applicable to other dialects found in the country.

6.3. Implications

The study has shown that LA does not diverge from the overall picture of the Arabic vocalic system. The number of vowels found in other Arabic dialects is also found in LA, which is 8 vowels in spite of the varying qualitative differences they exhibit from one dialect to another. The triangular shape that has often been claimed for the Arabic vowel system is obvious from the distribution of LA vowels in the vowel space, with the three fundamental vowels /iː, uː, aː/ occupying the three corners of this triangularly shaped vocalic system. Another similarity that LA shares with other Arabic dialects is the tendency of the short vowels to be more central in the vowel space while long vowels tend to be peripheral. However, this tendency is not uniform in all Arabic dialects, especially in the case of the low long vowel and its short counterpart. In some dialects, the short vowel tends to be more to the front in the vowel space (e.g. Iraqi) or lower (e.g. Cairene) or both (e.g. Jordanian) than its long counterpart (see table 4.4. in Chapter 4). Duration also exhibits a similar pattern in Arabic dialects. Similar ratios between long and short vowels are found in some Arabic dialects (e.g. Egyptian 0.40, Libyan 0.41, Sudanese 0.41) although the duration itself varies considerably from one dialect to another, which might be attributed to factors resulting from differences in methodological procedures (see Table 4.20 in Chapter 4).

In spite of these similarities shared by Arabic dialects, arriving at conclusions as to what these dialects have in common and what acoustic characteristics the Arabic language possesses still needs further research. One prospective area of such research is MSA. This is the variety that all speakers of Arabic dialects share, and by studying this variety as it is spoken by speakers of different dialects, more evidence for some of the previously-mentioned findings could be gathered so that the whole picture might become clearer.

Based on the qualitative difference found in the production of participants between short and long vowels, it was suggested earlier that different symbols to mark this qualitative difference should be used. However, the decision to use different symbols seems to be rather complicated. As mentioned earlier, one source of difficulty here is the results obtained from perception experiments which showed that this qualitative difference is less important in perception. Another factor is the discrepancy of results obtained from studies of other Arabic dialects. These dialects show a variety of possibilities regarding vowel qualities, ranging from the one that assumes qualitative equality between short and long vowels to those that show that short vowels are more peripheral than long vowels or, alternatively, that short vowels are more centralised while long vowels are marginal. As stated earlier, this diversity among Arabic dialects might have resulted from the differences in the methods adopted to obtain the data from one study to another.

Therefore, regarding the issue of using different symbols for short vowels from those used for long ones, more than one possibility exists. First, since the qualitative difference from which these symbols are derived is not important in perception, one might argue that the use of different symbols is not necessary. These vowels are produced to be perceived by listeners and, therefore, what counts is perception. Another argument supporting this is that the use of the same symbols is more convenient when cross-dialectal comparisons are made. For example, if LA is to be compared with other Arabic dialects, it would be more appropriate if the same symbols were used. However, production is as important as perception and the qualitative difference upon which the use of different symbols is based was found to be statistically significant. Therefore, underestimating it would be unreasonable.

The fact that the LA short low vowel tends to be higher and more retracted in the vowel space than its long counterpart leads to suggesting the use of /ə/ to refer to this short vowel. However, it should be emphasised that this vowel is not equal to the IPA 'schwa' which is higher and more retracted in the vowel space. Therefore, symbols used to represent vowels should be allowed to accommodate more variation in the vowel space rather than representing single dots. Productions of LA vowels showed variations represented by clouds in the vowel space, so representing them as dots rather than larger areas might be misleading.

In this study, listeners were found to rely heavily on durational cues in their perception of a short/long vowel contrast. In English it has been found (e.g. by Escudero 2000) that native speakers exclusively depend on spectral cues in distinguishing a tense/ lax contrast. On the other hand, Spanish listeners, who do not have durational contrast in their language, relied mainly on durational cues in contrasting English tense and lax vowels because durational cues are more prominent than spectral cues. Therefore, it is expected that LA speakers will do the same when learning to perceive English tense-lax contrasts, for example. However, a plausible question is whether they will also be able to learn to speak this contrast relying on durational or spectral differences, or both, since both types of cues are found to be significant in the production of vowels in LA.

Finally, it was found that production and perception differ considerably especially with regard to the qualitative difference between short and long vowels, which was found to be significant in production but not in perception. A plausible implication is that perception and production do not always go hand in hand and it is possible that "changes in production occur first, or that they occur in the absence of corresponding changes in perception" (Flege 1997, p.467). It would be expected, therefore, that a time would come when the qualitative difference realised in production might surface in perception in the future.

6.4. Limitations of the study

Although this study aimed at the investigation of the vocalic system of Libyan Arabic, it was limited to the examination of monophthongs. The diphthongs and vowels in emphatic context were excluded due to time and space limitations. However, a separate study (Kriba, forthcoming) is currently dealing with emphasis in LA, and emphatic vowels constitute a part of the focus of that study.

Another limitation is related to the sample. It might have been better if the speakers who participated in the perception experiment had also participated in the production experiment. However, due to the temporal gap between the two experiments it was not possible to reconstitute the same sample. However, efforts were made to control for other variables, including age, gender, dialect and foreign language experience.

A third limitation of the study concerns the exclusion of $f0$ from the analysis of vowel quality. A considerable literature has shown the effect of $f0$ on vowel quality. For example, Miller (1952) showed that raising the $f0$ of a synthesised vowel results in a shift of the vowel boundaries to higher frequencies by listeners. Kasuya et al (1968) and Hirahara et al (1992) also confirmed that there is a correlation between $f0$ and the other formant frequencies, in the sense that when $f0$ is higher, these other formants also become higher. However, in the perception experiment the $f0$ of the synthesised vowels was manipulated along with F1, F2 and F3 in order to make the stimuli sound more natural.

The perception experiment was limited to the investigation of the two high vowels. This investigation has to some extent clarified the perception of these two vowels and helped in making a connection between speech perception and speech production. However, it could have been much better if all the vowels investigated in production were also examined in perception. This could have allowed us to discover to what extent the production of the whole vocalic system matches the perception of that system. This could have enabled a test of the hypothesis proposed by Johnson, Flemming & Wright (1993) that the perceptual vocalic space corresponds to a hyper-articulated space, which has been found to be true in English (Evans & Iverson 2002) and other Arabic dialects (Al-Tamimi et al 2002) . However, the results of the perception experiment do give some hint of the validity of the hyperspace hypothesis. Stimuli at the endpoints of all continua used in the experiment were perceived as belonging to one category or the other by 100% of participants. This percentage is lower towards the middle of these continua and would probably remain high if these continua were extended outwards with higher values of formants. The assumption is that if these endpoints were extended then there would be agreement among all listeners.

6.5. Suggestions for further research

LA phonology is still a relatively understudied area of investigation. The number of studies conducted with a focus on LA is very small, and some aspects of the dialect still need to be investigated. Therefore, research is encouraged to explore this dialect and its characteristics. Although this study has aimed at providing a detailed description of the vocalic system of the dialect, there remain some aspects that need to be covered in future research.

Firstly, it was found that Libyan short and long vowels differ not only in their quantity but also in their quality. Similar findings have also been reported in other Arabic dialects (see section 4.2.1 in Chapter Four). This is not in harmony with what is generally thought to be found in the standard variety of Arabic known as MSA. Some researchers (e.g. Alghamdi 1998) have claimed that MSA vowels are shaped by dialectal realisations of these vowels. However, this claim is based on studies in which MSA words were produced by participants from different dialectal backgrounds. Since a considerable number of words exist in both the dialect and the standard variety, it would be interesting if a study was carried out to compare words as produced in the dialect with the same words as realised by the same speakers in MSA.

LA diphthongs have not been dealt with acoustically, so one of the suggestions for future research is to study these diphthongs. Arabic diphthongs have been described as having different temporal organisation from their counterparts in English, in that their first element is shorter than their second element (Muftah 2001). However, this view is based on impressionistic auditory examination. Therefore, an experimental investigation is required.

Another prospective area of research is the perception of these vowels by native speakers. This study has only covered two vowels, and even with these two vowels not all possibilities were investigated. It is suggested that future research investigates the perception of the other two sets of short and long vowels in order to find out whether the reliance on duration in distinguishing between /iː/ and /ɪ/ also holds for the other pairs.

It was also found in the results of the perception experiment that listeners tend to be more certain about their choice of stimuli at or near the end-points of the continua, and the further the distance from these endpoints the less certain listeners become about the identity of the vowel. This might give some support to the assumption made by Johnson et al (1993) and Johnson (2000), who maintain that the perceived vowel space corresponds to a hyper-articulation space, and this was found to be true in Jordanian and Moroccan Arabic (Al-Tamimi et al 2002). Therefore, a perception study which includes all vowels in the vowel space will give more evidence for this assumption and the way a system of vowels functions as a whole.

Finally, another suggested study could be one that aims to test listeners' choices of the best prototype of a vowel category. One of the aims of the perception experiment conducted in this study was to specify the boundary between vowel categories. Therefore, a forced choice technique was used and listeners had to say whether the stimulus they heard was this or that vowel. A suggested experiment would allow participants to choose from all stimuli belonging to a vowel category the one that they think best represents the vowel type (known as the method of adjustment (MOA), Johnson et al 1993). A comparison of their choices would determine whether or not they vary in their choice of the best prototype and to what extent. Then these prototypes could be compared to the participants' produced vowels to find out to what extent their perception matches their production of these vowels. This method has been widely used by researchers working on perception (see, for example, Frieda et al 2000).

A final suggested work is related to the generalisation of the findings of this study. As stated in the introduction to this study, Libya is a very large country and there is a variety of dialects which differ considerably, especially at the suprasegmental level. However, these dialects bear a segmental similarity as the review of literature has shown. The vowels investigated in this study do not only constitute the vocalic system of most of the dialects found across the country but also those available in other Arabic dialects in spite of the slight variation in their realisations from one Arabic dialect to another. In order for the results of this study to be generalised, other studies related to other Libyan dialects should be conducted. A cross-dialect comparison will reveal whether these vowels, though available in all Libyan dialects, are realised similarly or if, and to what extent, they differ.

REFERENCES

Abdul-Jaleel, A., 1998. *ʔal-ʔaswa:t ʔal-luɣawiyah.* Amman, Jordan: Dar Safa Publishing & Distributing.

Abercrombie, D., 1967. *Elements of general phonetics.* Chicago: Aldine.

Abou Haidar, L., 1994. 'Norme linguistique et variabilité dialectale: analyse formantique du système vocalique de la langue arabe', *Revue de Phonétique Appliquée*, 110, pp. 1-15.

Abramson, A., and Lisker, L., 1967. Discriminability along the voicing continuum: Cross-language tests. *Proceedings of the Sixth International Congress of Phonetic Sciences.* Prague: Academia Publishing House of the Czechoslovak Academy of Sciences.

Abumdas, A., 1985. *Libyan Arabic phonology.* Thesis (PhD). University of Michigan.

Al-Ageli, H. M., 1995. *Syllable and metrical structure in Tripolitanian Arabic.* Thesis (PhD). University of Essex.

Al-Ani, S. H., 1970. *Arabic phonology; an acoustical and physiological.* The Hague, Mouton.

Alghamdi, M., 1998. A spectrographic analysis of Arabic vowels: A cross-dialect study. *Journal of King Saud University*: Arts, 10 (1), 3-24.

Alghamdi, M., 1990. *Analysis, synthesis and perception of voicing in Arabic.* Thesis (PhD). Reading University.

Al-Masri, M. & Jongman, A., 2004. Acoustic correlates of emphasis in Jordanian Arabic: Preliminary Results. In A. Agwuele et al., ed. *Proceedings of the 2003 Texas Linguistic Conference: Coarticulation in speech production and perception.* Somerville, MA: Cascadilla Proceedings Project, 96-106.

Al-Tamimi, J., Barkat-Defradas, M., 2002. Inter-dialectal and inter-individual variability in production and perception: A preliminary study in Jordanian and Moroccan Arabic. [online] Available from: http://www.ddl.ish-lyon.cnrs.fr/fulltext/Al-Tamimi/Al-Tamimi_2002_AIDA.pdf [25 September 2006].

Al-Tamimi, J., Carre, R. & Marsico, E., 2004. The status of vowels in Jordanian and Moroccan Arabic: Insights from production and perception. [online] Available from: http://www.ddl.ish-lyon.cnrs.fr/fulltext/Al-Tamimi/Al-Tamimi_2004_Poster_ASA.pdf [25 September 2006].

Al-Tamimi, J. & Ferragne, E., 2005 Does vowel space size depend on language vowel inventories? Evidence from two Arabic dialects and French. [online] Available from: http://www.ddl.ish-lyon.cnrs.fr/fulltext/Al-Tamimi/Al-Tamimi_2005_Interspeech.pdf [23 December 2007].

Arai, T., 2004. Formant shift in nasalization of vowels. *Journal of the Acoustical Society of America*, 115 (5), p. 2541.

Aurayieth, A., 1982. *The phonology of the verb in Libyan Arabic.* Thesis (PhD). University of Washington

Aylett, M., Building a statistical model of the vowel space for phoneticians. [online] Available from: http://www.cogsci.ed.ac.uk/~matthewa/publications/icslp98c.pdf [06 July 2008].

Barkat-Defradas, M., Al-Tamimi, J & Benkirane, T., 2003. Phonetic variation in production and perception of speech: a comparative study of two Arabic dialects. [online] Available from: http://www.ddl.ish-lyon.cnrs.fr/fulltext/AL-Tamimi/Barkat-Defradas_2003_ICPhS.pdf [24 October 2007].

Beckman, M.E., Jung, T. P., Lee, S., de Jong, K., Krishnamurthy, A. K., Ahalt, S. C., & Cohen, K.B., 1995. Variability in the production of quantal vowels revisited. *Journal of the Acoustical Society of America*, 97, 471-90.

Belkaid, Y., 1984. Les voyelles de l'arabe littraire modern. Analyse Spectrographique. *Travaux de l'Institut de Phonetique de Strasbourg*, 16, 217-240.

Best, C. T., 1995. A direct realist view of cross-language speech perception. *In* W. Strange, ed. *Speech Perception and Linguistic Experience* (pp. 171-206). Baltimore, MD: York Press.

Bin-Muqbil, M. S., 2006. *Phonetic and phonological aspects of Arabic emphatics and gutturals.* Thesis (PhD). University of Wisconsin-Madison.

Bladon, A., 1985. Diphthongs: A case study of dynamic auditory processing. *Speech communication*, 4, 145-54.

Blumstein, S. E., & Stevens, K. N., 1980. Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America*, 66, 1001-1017.

Boersma, P., & Weenink, D. Praat: doing phonetics by computer. [online] Available from: http://uvafon.hum.uva.nl/praat/ [21 October 2006].

Bohn, O. S., 1995. Cross-language speech perception in adults: First language transfer doesn't tell it all. *In* W. Strange, ed. *Speech perception and linguistic experience: Issues in cross-language research.* Timonium, MD: York Press, 279-304.

Botagga, M. F., 1991. *A generative phonetic feature analysis of the vowel development of native Arabic speakers learning English as a second langaguage.* Thesis (PhD). Wales, Cardiff University

Bryman, A., 2001. *Social research methods.* Oxford: Oxford University Press.

Bukshaisha, F., 1985. *An experimental phonetic study of some aspects of Qatari Arabic.* Thesis (PhD). Edinburgh University.

Byrd, D.,1994. Relations of sex and dialect to reduction. *Speech Communication 15*, 39-54.

Card, E. A., 1983. *A phonetic and phonological study of Arabic emphasis*, Thesis (PhD). Cornell University.

Carre, R. & Mrayati, M., 1990. Articulatory-acoustic-phonetic relations and modelling, regions and modes. *In* W. J. Hardcastle & A. Marchal, eds. *Speech production and speech modelling.* Netherlands: Kluwer Academic Publishers, 211-240.

Chomsky, N. & Miller, G. A., 1963. Introduction to the formal analysis of natural Languages. *In* Luce R. Duncan, Robert R. Bush and Eugene Galanter, eds. *Handbook of mathematical psychology.* New York: Wiley, 269-321.

Chomsky, N. & Halle, M., 1968. *The sound pattern of English.* New York: Harper and Row.

Clements, G. N., 2003. Feature economy in sound systems. *Phonology* 20 (3), 287-333.

Clermont, F., 2007. Formant-pattern estimation guided by cepstral compatibility. Abstract Proc. Int. Conf. on Forensic Phonetics and Acoustics, Plymouth, 22-25 July 2007.

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. *Journal of the Acoustical Society of America*, 24, 597-606.

Cowan, W., 1970. The vowels of Egyptian Arabic. *Word*, 26, 94-100.

Daniloff, R. G., & Hammarberg, R. E., 1973. On defining coarticulation. *Journal of Phonetics*, 1, 239-48.

Davis, S., 1991. On the phonological analysis of emphasis in two modern Arabic dialects. *International journal of Islamic and Arabic studies*, 8 (2), 1-20.

Davis, S., 1993. Arabic pharyngealization and phonological features. *In.* M. Eid & C. Holes, eds., *Perspectives on Arabic Linguistics V.* Amsterdam: John Benjamins, 149-162.

Davis, S., 1995. Emphasis spread in Arabic and grounded phonology. *Linguistic inquiry*, 26 (3), 465-476.

Delattre, P. C., Liberman, A. M., & Cooper, F. S & Gerstman, L. J., 1952. An experimental study of the acoustic determinants of vowel colour: Observations on one- and two-formant vowels synthesized from spectrographic patterns. *Word*, 8, 195-210.

Delattre, P. C., Liberman, A. M., & Cooper, F. S.,1955. Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, 27, 769-73.

Delattre, P., 1971, Pharyngeal features in the consonants of Arabic, German, Spanish, French, and American English. *Phonetica*, 23, 129-155.

Denes, P. B. & Pinson, E. N., 1993.*The speech chain: The physics and biology of spoken language*. New York: W. H. Freeman and Company.

Denscombe, M., 2003. *The good research guide*. Maidenhead: Open University Press

Dickens, J., 1996. Some speculations on the vowel system of classical Arabic. *In* T. Niblock, A. Ehteshami, P. Starkey & F. Faqir, eds. *Three topics in Arabic phonology.* Centre for Middle Eastern and Islamic Studies. University of Durham, 6-19.

Disner, S. F., 1986. On describing vowel quality. *In* J. J. Ohala & J. J. Jaeger, eds. *Experimental phonology*. Academic Press inc., 69-79.

Docherty, G. and Foulkes, P., 2000. Speaker, speech, and knowledge of sounds. *In* N. Burton-Roberts, P. Carr, and G. Docherty, eds. *Phonological knowledge: conceptual and empirical issues.* Oxford: Oxford University Press, 105-30.

Elgadi, A., 1986. *Tripolitanian Arabic phonology and morphology: A generative approach*. Thesis (PhD). Georgetown University.

Escudero, P., 2000. *Developmental patterns in the adult L2 acquisition of new contrasts: The acoustic cue weighting in the perception of Scottish tense/lax vowels by Spanish speakers*. Thesis (M.Sc.). University of Edinburgh.

Escudero, P., 2002.The perception of English vowel contrasts: Acoustic cue reliance in the development of new contrasts. *In* Jonathan Leather & Allan James, eds. *Proceedings of the 4th Symposium on the acquisition of second language speech, New Sounds 2000,* pp. 122-31. University of Klagenfurt.

Evans, B. G., & Iverson, P., 2002. Vowel Normalisation for dialects. In S. Hawkins and N. Nguyen, eds. *Temporal Integration in the Perception of Speech*, p.35. Cambridge, UK: Centre for Research in the Arts, Social Sciences, and Humanities.

Fant, G. 1958. Modern instruments and methods for acoustic studies of speech. *Proceedings of the 8th International Congress of Linguistics*. Oslo University Press, 282-358.

Fant, G., 1960. *Acoustic theory of speech production*. The Hague: Mouton.

Fant, G., 1962. Descriptive analysis of the acoustic aspects of speech. *Logos*, 5, 3-17.

Fant, G., 1972. *Vocal tract wall effects, losses, and resonance bandwidths*. STL-QPSR, 2-3, 28-52.

Fant, G., 1973. *Speech sounds and features*. Cambridge, MA: MIT Press.

Ferguson, C., 1956. The emphatic l in Arabic. *Language*, 32 (2), 446-452.

Ferguson, C., 1959. Diaglossia. *Word*, 15, 325-40.

Flege, J., 1979. *Phonetic interference in second language acquisition*. Thesis (PhD). Indiana University.

Flege, J. & Port, R., 1981. Cross-language phonetic interference: Arabic to English. *Language and Speech*, 24, 125-46.

Flege, J., 1987. The production of new and similar phones in a foreign language: Evidence for the Effect of Equivalence Classification. *Journal of Phonetics*, 15 (1), 47-65.

Flege, J. & Eefting, W., 1987. Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6, 185-202.

Flege, J., 1997. Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics* 25, 437-70.

Fodor, J. A., 1983. *The modularity of mind*. Cambridge, MA: MIT Press.

Fowler, C. A., 1986. An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14, 3-28.

Fowler, C. A. & Galantucci, B., 2005. The relation of speech perception and speech production. *In* D. B. Bison & R. E. Remez, eds. *The Handbook of Speech Perception*. Blackwell Publishing, 633-52.

Frieda, B. M., Walley, A. C. & Sloane, M. E., 2000. Adults' perception and production of the English vowel /i/. *Journal of Speech, Language, and Hearing Research*, 43, 129-43.

Fry, D. B., Abramson, A. S., Eimas, P. D., & Liberman, A. M., 1962. The identification and discrimination of synthetic vowels. *Language and Speech*, *5*, 171-189.

Fujisaki, H. & Kawashima, T., 1968. The role of pitch and higher formants in the perception of vowels. *IEEE Audio Electronics*, AU-16 (1) 73-7.

Gairdner, W. H. T., 1925. *The phonetics of Arabic*. London: Oxford University Press.

Gay, T., 1978. Effect of speaking rate on vowel formant movements. *Journal of the Acoustical society of America*, 63, 223-30.

Ghazeli, S., 1977. *Back consonants and backing coarticulation in Arabic*. Thesis (PhD). The University of Texas at Austin

Giegerich, H., 1992. *English phonology: An introduction*. Cambridge University Press: Cambridge.

Goldinger, S. D., 1998. Echoes of echoes: An episodic theory of lexical access. *Psychological Review*, 105, 251–279.

Halberstam, B.,1998. *Vowel normalization: The role of fundamental frequency and upper formants*. Thesis (PhD).City University of New York.

Harrell, R. S., 1957. *The phonology of colloquial Egyptian Arabic*. New York: American Council of Learned Societies.

Harrison, P., 2004. *Variability in formant measurements*. Thesis (MA). University of York, UK.

Harrington, J. & Cassidy, S., 1994. Dynamic and target theories of vowel classification: Evidence from monophthongs and diphthongs in Australian English. *Language and Speech* 37 (4), 357-73.

Hassan, Z, M., 1981. *An experimental study of the vowel duration in Iraqi spoken Arabic*. Thesis (PhD). UK: Dept of Linguistics & Phonetics, University of Leeds.

Hassan, Z, M., 2002. Gemination in Sweedish & Arabic with a particular reference to the preceding vowel duration. An instrumental & comparative approach. *TMH-QPSR*, 44, 81-4.

Hassan, Z, M., 2003. Temporal compensation between vowel and consonant in Swedish & Arabic in sequences of CV:C & CVC: and the word overall duration . *PHONUM*, 9, 45-48.

Hawkins, S., 2004. Puzzles and patterns in 50 years of research on speech perception. *From Sound to Sense: 50+ Years of Discoveries in Speech Communication*, MIT, Cambridge, MA, USA http://www.rle.mit.edu/soundtosense/conference/pdfs/invitedspeakers/Hawkins%20PAPER.pdf.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K., 1995. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America.* 5, 3099-3111.

Hirahara, T. & Kato, H., 1992.The effect of F0 on vowel identification. *In* Y. Tohkura, E. Vatikiotis-Bateson & Y. Sagisaka, eds. *Speech perception, production and linguistic structure.* Tokyo: IOS Press. pp. 89-112.

Hockett, C., 1955. A manual of phonology. *International Journal of American Linguistics* **21**(4): Part 1. Memoir 11

Hombert, J. M., & Puech, G.,1984. Espace vocalique et structuration perceptuelle : application au Swahili, *Pholia* 1 :199-208. http://www.ddl.ish-lyon.cnrs.fr/fulltext/gabon/pholiapdf/pholia1/Hombert_swahili.pdf

Hoopingarner, D., 2004. *Native and nonnative differences in the perception and production of vowels.* Thesis (PhD). Michigan State University.

Huang, C. B., 1992. Modelling human vowel identification using aspects of formal trajectory and context. *In* Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka, eds. *Speech perception, production and linguistic structure* (IOS, Amsterdam), pp. 43-61.

Hussain, A., 1985. *An experimental investigation of some aspects of the sound system of the Gulf Arabic dialect with special reference to duration.* Thesis (PhD) University of Essex.

Huthaily, K., 2003. *Contrastive phonological analysis of Arabic and English.* Thesis (MA). The University of Montana.

Janson, T., 1986. Sound change in perception: An experiment. *In* J. J. Ohala & J. J. Jaeger, eds. *Experimental Phonology.* Academic Press inc., 253-260.

Johnson, K., Flemming, E. & Wright, R., 1993. The hyper space effect: phonetic targets are hyperarticulated, *Language.* 69, 505-528.

Johnson, K., Ladefoged, P. & Lindau, M., 1993. Individual differences in vowel production. *Journal of the Acoustical Society of America. 94*, 701-714.

Johnson, K., 2005. Speaker normalisation in speech perception. *In* D. Pisoni & R. Remez. *The handbook of Speech Perception.* Blackwell Publishing Ltd., 363-389.

Johnson, K., 2003. *Acoustic and auditory phonetics.* Oxford: Blackwell Publishing.

Johnson, K., 2000. Adaptive dispersion in vowel perception. *Phonetica* 57, 181-88.

Jones, D., 1964. *An Outline of English Phonetics,* W. Heffer & Sons, Cambridge, ninth edition.

Jones, D., 2003. *English pronouncing dictionary.* 16th ed. Edited by Roach, P., Hartman, J. & Setter, J. Cambridge, U.K. : Cambridge University Press

Joos, Martin., 1948. Acoustic phonetics. *Language Monographs 23.* Baltimore: Linguistic Society of America.

Kasuya, H., Suzuki, H. & Kido, K., 1968. Changes in pitch and first three formant frequencies of five Japanese vowels with age and sex of speakers. *Journal of the Acoustical society of Japan*, 24 (6), 355-64.

Keating, P. A., Mikos, M. J. & Ganog, W. F., 1981. Across-language study of range of voice onset time in the perception of initial stop voicing. *Journal of the Acoustical Society of America,* 70, 1261-1271.

Kriba, H. (forthcoming). *Acoustic parameters of emphasis in Libyan Arabic.* Unpublished PhD Thesis, Newcastle University, UK.

Klatt, D. H., 1976. The linguistic uses of segmental duration in English: acoustic and perceptual evidence. *Journal of the Acoustical society of America*, 59, 1208-21.

Klatt, D. H., 1988. Review of selected models of speech perception. *In* W. Marslen Wilson, ed. *Lexical representation and process.* Cambridge: MIT Press, 169-226.

Klein, W., Plomp, R., & Pols, L. C. W., 1970. Vowel spectra, vowel spaces, and vowel identification. *Journal of the Acoustical society of America*, 48, 999-1009.

Kuhl, P. K., 1991. Human adults and human infants show a perceptual magnet effect for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, 50, 93-107.

Kuhl, P. K, 2000. Language, mind and brain: Experience alters perception. *In* M. S. Gazzaniga, ed. *The new cognitive neurosciences* (2nd ed.). Cambridge, MA: MIT Press, 99-115.

Kuhl, P. K. & Iverson, P., 1995. Linguistic experience and the perceptual magnet effect. *In* W. Strange, ed. *Speech perception and linguistic experience: issues in cross-language research.* Baltimore: York Press, 121-54.

Ladefoged, P. & Broadbent, D. E., 1957. Information conveyed by vowels. *Journal of the Acoustical Society of America*, 29, 98-104.

Ladefoged, P., 1967. *Three areas of experimental phonetics.* Oxford: Oxford University Press.

Ladefoged, P., 1996. *Elements of acoustic phonetics.* Chicago and London: the University of Chicago Press.

Ladefoged, P., 2001. *Vowels and consonants: an introduction to the sounds of language.* Malden, Mass.: Blackwell, 2001.

Ladefoged, P., 2003. *Phonetic data analysis: an introduction to field work and instrumental techniques*. USA: Blackwell Pub.

Ladefoged, P. & Maddieson, I., 1996. *The sounds of the world's languages*. Oxford, Blackwell.

Lang, E & Ohala, J., 1996. Temporal cues for vowels and universals of vowel inventories. [online] Available from: http://linguistics.berkeley.edu/phonlab/users/ohala/papers/lang_and_ohala_icslp_1996.pdf [05 Jan 2007].

Laradi, W. J., 1983. *Pharyngealization in Libyan (Tripoli) Arabic: an instrumental study*. Thesis (PhD). University of Edinburgh.

Lehiste, I. & Peterson, G., 1961. Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America*, 33, 419-425.

Lehiste, I.,1970. *Suprasegmentals*. Cambridge: MIT Press.

Liberman, A M, Harris K S, Hoffman H S, Griffith B. C., 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of experimental Psychology* 54: 358-68, 753-71.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M., 1967. Perception of the speech code. *Psychological Review*, 74, 431-61.

Liberman, A. M., 1982. On finding that speech is special. *American Psychologist*, 37, 148-167.

Liberman, A. M., & Mattingly, I. G., 1985. The motor theory of speech perception revised. *Cognition*, 21, 1-36.

Lindblad, P., 2001. Critique of the IPA vowel quadrangle, especially the use of vowel points in it. Lund University, Dept. of Linguistics. *Working Papers*, 49, 104-107.

Lindblom, B. E. F., 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 35, 1773-81.

Lindblom, B., 1967. Vowel duration and a model of lip mandible coordination. *Quaterly Progress & Staus Report Speech Transmission Laboratory*, pp. 1-29.

Lindblom, B., 1990. Explaining phonetic variation: A sketch of the H & H theory. *In* W. Hardcastle & A. Marchal, eds. *Speech production and speech modelling*. Netherlands: Kluwer Academic Publishers. 403-39.

Lindblom, B., 1996. Role of articulation in speech perception: Clues from production. *Journal of the Acoustical society of America*, 99 (3), 1683-92.

Lindblom, B. & Studdert-Kennedy, M., 1967. On the role of formant transitions in vowel recognition. *Journal of the Acoustical Society of America*, 42 (4), 830-43.

Lofqvist, A., 1990. Speech as audible gestures. *In* W. J. Hardcastle & A. Marchal, eds. *Speech production and speech modelling*. Netherlands: Kluwer Academic Publishers. 289-322.

Maddieson, I., 1984. *Patterns of Sounds. Cambridge studies in speech science and communication*. Cambridge: Cambridge University Press.

MacNeilage, P. F., 1970. Motor control of serial ordering of speech, *Psychological Review*. 77, 182-96.

Makashay, M.J., 2003. *Individual differences in speech and non-speech perception of frequency and duration*. Thesis (PhD). Ohio State University.

Massaro, D. W., & Cohen, M. M., 1983. Evaluation and integration of visual and Auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 753-71.

Massaro, D. W., 1989. Multiple book review of speech perception by ear and eye: A paradigm for psychological inquiry. *The Behavioral and Brain Sciences*, 12, 741-55.

Massaro, D. W., 1992. The fuzzy Logical model of speech perception: A framework for research and theory. *In* Y. Tohkura, E. Vatikiotis & Y. Sagisaka, eds. *Speech perception and linguistic experience*. Japan: ATR Auditory and Visual Perception Research Laboratories, 79-82.

Massaro, D. W., 2001. Speech perception. *In* N. M. Smelser & P.B. Baltes & W. Kintsch, eds. *International encyclopedia of social and bahavioral sciences*. Amsterdam: Elsevier. 14870-14875.

Mattingly, I. G. & Liberman, A. M., 1969. The speech code and the physiology of language. *In* K. N. Leibovic. ed. *Information processing in the nervous system*. New York: Springer- Verlag.

McGurk, H. & MacDonald, J., 1976. Hearing lips and seeing voices. *Nature*, 264, 746-748.

Miller, R. L., 1952. Auditory tests with synthetic vowels. *Journal of the Acoustical Society of America*, 25, (1) 114-21.

Mitchell, T. F., 1993. *Pronouncing Arabic*. 2^{nd} vol., Oxford University Press.

Mitleb, F. M., 1984a. Voicing effect on vowel duration is not an absolute universal. *Journal of Phonetics*, 12, 23-27.

Mitleb, F. M., 1984b. Vowel length contrast in Arabic and English: A spectrographic test. *Journal of Phonetics*, 12, 229-35.

Mora, J. C. & Fullana, N., 2007. Production and perception of English /i:/-/ɪ/ and /æ/-/ʌ/ in a formal setting: Investigating the effect of experience and starting age.[online] Available from: http://www.icphs2007.de/conference/Papers/1594/1594.pdf [21 October 2007].

Motoki, K., 2002. Three-dimensional acoustic field in vocal tract. [online] Available from: http://www.jstage.jst.go.jp/article/ast/23/4/207/_pdf [11 December 2007].

Muftah, A., 2001. *English and Arabic sound systems: Similarities and differences and their effect on learning English by Arab students*. Thesis (MA). Tripoli: Al-Fateh University.

Nasr R. T., 1959. Velarization in Lebanese Arabic. *Phonetica,*3, 203-9.

Nearey, T. M., 1989. Static, dynamic, and relational properties in vowel perception. *Journal of the Acoustical Society of America*, 85, 2088-113.

Nearey, T. M.,1997. Speech perception as pattern recognition. *Journal of the Acoustical Society of America*, 101 (6), 3241-54.

Newman, D. L., 2002. The phonetic status of Arabic within the world's languages. *Antwerp Papers in Linguistics*, 100, 65-75.

Newman, D. L. & Verhoeven, J., 2002. Frequency analysis of Arabic in connected speech. *Antwerp Papers in Linguistics*, 100, 77-86.

Norlin, K., 1987. *A phonetic study of emphasis and vowels in Egyptian Arabic*. Thesis (PhD). Leads: University of Leeds.

Nosofsky, R.M., 1988. Exemplar-based accounts of relations between classification, recognition and typicality. *Journal of Experimental Psychology: Learning Memory, and Cognition 14*, 700-708.

Nusbaum, H. C. & Morin, T. M.,1992. Paying attention to differences among talkers. *In* Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka, eds. *Speech Perception, Production and Linguistic Structure*. Burke: IOS Press, 113-23.

Obrecht, D., 1968. *Effects of the second formant on the perception of velarization consonants in Arabic*. Paris: Mouton.

Oden, G. C., & Massaro, D. W., 1978. Integration of featural information in speech perception. *Psychological Review*, 85, 172-91.

Ohala, J. J., 1996. Speech perception is hearing sounds, not tongues. *Journal of the Acoustical Society of America*, 99 (3), 1718-25.

Perkell, J. S., 1990. Testing theories of speech production: implications of some detailed analysis of variable articulatory data. *In* W. J. Hardcastle & A. Marchal, eds. *Speech production and speech Modelling*. Dordrecht: Kluwer Academic Publishers, 263-83.

Peterson, G.E., 1951. The phonetic value of vowels. *Language* 27 (3), 541-53.

Petrson, G. E. & Barney, H. L., 1952. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 24 (2), 175-84.

Petrson, G. E. & Lehiste, I.,1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, 32 (6), 693-703.

Pisoni, D. B.,1997. Some thoughts on normalization in speech perception. *In* K. Johnson & J.W. Mullennix, eds. *Talker variability in speech processing*. San Diego: Academic Press, 9-32.

Pisoni, D. B., 1973. Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, 13 (2), 253-260.

Pisoni, David B.,1971. *On the nature of categorical perception of speech sounds.* Thesis (PhD).University of Michigan.

Pisoni D. B. & Lively S. E. 1995.Variability and Invariance in Speech Perception: A New Look at Some Old Problems in Perceptual Learning. *In* W. Strange, ed.: *Speech perception and linguistic experience: issues in cross language research.* Baltimore: York Press.

Port, R., Al-Ani, S. & Maeda, S., 1980. Temporal compensation and universal phonetics. *Phonetica*, 37, 235-252.

Punch, K., 2000. *Developing effective research proposal.* London: SAGE.

Raphael, L. J., 2005. Acoustic cues to the perception of segmental phonemes. *In* D. B. Bison & R. E. Remez, eds. *The Handbook of Speech Perception.* Blackwell Publishing,183-206.

Repp, B., Healy, A. F. & Crowder, R. G., 1979. Categories and context in the perception of isolated steady-state vowels. *Journal of Experimental Psychology: Human Perception and Performance,* 5, 129-145.

Roca, I. & Johnson, W., 1999. *A course in phonology.* Oxford: Blackwell Publishers.

Rosch, E., 1975. Cognitive reference points. *Cognitive Psychology*, 7, 532-547.

Rosner, B. S. & Pickering J. B., 1994. *Vowel perception and production.* Oxford: Oxford University Press.

Ryalls, J., 1996. *A basic introduction to speech perception.* London: Singular Puplishing Group, Inc.

Schwartz, J., Boe, L., Vallee, N. & Abry, C., 1997. Major trends in vowel system inventories. *Journal of phonetics*, 25, 233-253.

Skinner, T. E.,1977. Speaker invariant characteristics of vowels, liquids, and glides using relative formant frequencies. *Journal of the Acoustical Society of America*, 62 (S1), S5.

Smith, R., 2004. *The role of fine phonetic detail in word segmentation.* Thesis (PhD). Department of Linguistics, Cambridge University.

Sonderegger, M. A., 2004. *Subglottal coupling and vowel space: An investigation in quantal theory.* Thesis (M.Sc). Massachusetts Institute of Technology.

Stephenson, L., 2004. Lexical frequency and neighbourhood density effects on vowel production in words and nonwords. *Proceedings of the 10th Australian International Conference on Speech Science & Technology.* Macquarie University, Sydney, December 8 to 10, 2004. Australian Speech Science & Technology Association Inc.

Stevens, K. N., 1972. The quantal nature of speech: Evidence from articulatory acoustic data. *In* E. E. David, Jr., and P. B. Denes, eds. *Human communication: A unified view.* New York: McGraw-Hill, 51-66.

Stevens, K. N., 1989. On the quantal nature of speech. *In* R. D. Kent, B. S. Atal and J. L. Miller, eds. *Papers in speech communication: Speech production.* USA: Journal of the Acoustical Society of America, 357-99.

Stevens, K. N., & House, A. S., 1963. Perturbation of vowel articulations by consonantal context: An acoustic study. *Journal of Speech and Hearing Research*, 6, 111-28.

Stork, F. C. & Widdoson , J. D. A., 1974. *Learning about linguistics.* London: Hutchinson.

Strange, W., 1989. Evolving theories of vowel perception. *Journal of the Acoustical Society of America,* 85 (5), 2081-87.

Strange, W., Jenkins, J. J., & Johnson, T. L., 1983. Dynamic specification of coarticulated vowels. *Journal of the Acoustical Society of America*, 74 (3), 695-705.

Studdert-Kennedy, M., 1976. Speech perception. *In* N. J. Lass, ed. *Contemporary issues in experimental linguistics.* New York: Academic Press, 213-293.

Studdert-Kennedy, M., 1993. Discovering phonetic function. *Journal of Phonetics*, *21*, 147–155.

Tatham, M. & Morton K., 2006. *Speech production and perception.* New York: Palgrave MacMillan.

Traunmüller, H., 1990. Analytical expressions for the tonotopic sensory scale. *Journal of the Acoustical Society of America,* 88(1), 97-100.

Traummuller, H., 2007. Coarticulatory effects of consonants on vowels and their reflection in perception. [online] Available from: http://www.ling.su.se/staff/hartmut/reattri.pdf [20 October 2007].

Turk, A., Nakai, S. & Sugahara, M., 2006. Acoustic segment durational in prosodic research: A practical guide. [online]. Available from: http://www.let.uu.nl/LOT/GraduateProgram/LotSchools/Summerschool2005/course%20descriptions/Remeijsen_TurkEtal_Bert.pdf [5 July 2007].

Umeda, N., 1977. Consonant duration in American English. *The Journal of the Acoustical Society of America*, 61 (3), 846-58.

Van Bergem, D., 1995. *Acoustic and lexical vowel reduction.* Thesis (PhD). University of Amsterdam.

Van Bergem, D., 1994. A model of coarticulatory effects on the schwa. Speech *Communication,* 14: 143-162.

Wardhaugh, R., 2002. *An introduction to sociolinguistics.* 4th ed. USA: Blackwell Publishing

Watson, J. C. E., 1999. The directionality of emphasis spread in Arabic. *Linguistic Inquiry*, 30 (1), 289-300.

Watson, C. I. & Harrington, J., 1999. Acoustic evidence for dynamic formant trajectories in Australian English vowels. *The Journal of the Acoustical Society of America*, 106 (1), 458-68.

Watt, D. & Fabricius, A., 2003. Evaluation of a technique for improving the mapping of multiple speakers' vowel space in the F1~F2 Plane. [online]. Available from: http://www.york.ac.uk/sociolinguistics-archive/index.cfm?orderon=CLIENTFILEEXT&order=DESC&searchtype=Author&searchterms=&StartRow=1&PageNum=1 [1 June 2006].

Whalen, D. & Beddor, P., 1989. Connections between nasality and vowel duration and height: Elucidation of the Eastern Algonquian intrusive nasal. *Language*, 65, 457-486.

Willerman, R. & Kuhl, P., 1996. Cross language speech perception: Swedish, English, and Spanish speakers' perception of front rounded vowels. *In* H. T. Bunnell & W. Idsardi, eds. *Proceedings of ICSLP*, 96 (1) 442-45.

Wright, R., Frisch, S. & Pisoni, D. B., 1997. Research on spoken language processing. *Progress report No 21.* Bloomington, Indiana: Indiana University.

Younes, M., 1993. Emphasis spread in two Arabic dialects. *In*. M. Eid and C. Holes, eds., *Perspectives on Arabic linguistics V.* Amsterdam and Philadelphia: John Benjamins, 119- 145.

Younes, M., 1994. On emphasis and /r/ in Arabic. *In*. M. Eid, V. Cantarino and K. Walters, eds, *Perspectives on Arabic linguistics VI* Amsterdam and Philadelphia: John Benjamins, 215-235.

Youssef, I. M. 2006. *Vocalic feature assimilation in Cairene Arabic and Buchan Scots: Contrast and phonological activity in the parallel structure model of feature geometry.* Thesis (MA). I Tromsø University.

Zahorian, S. & Jagharghi, A., 1993. Spectral-shape features versus formants as acoustic correlates for vowels. *The Journal of the Acoustical Society of America,* 94, 1966-82.

Zawaydeh, B., 1999. *The phonetics and phonology of gutturals in Arabic.* Thesis (Ph D). Indiana University, Bloomington.

Zwicker, E., 1961. Subdivision of the audible frequency range into critical bands. *The Journal of the Acoustical Society of America*, 33:248–249, 1961.

Appendices

Appendix 1

LA words as used in production task

أنا قتلا   تَعالى   يا أخي

بِيك

بُوم

فِن

بُك

لَام

دُور

لِم

بِين

دِين

دُوم

بِن

بُل

سَاد

لُون

سَد

بِيت

بَرا

أنا قتلا   َتعالى   يا أخي

ِقيس

ُلوم

ِفز

ُفل

َفات

ُلوز

َفد

ِسيف

ِلين

ُموس

هِز

بُد

نَاب

ُفوز

َنق

ِسير

َبرا

أنا قتلا َتعالى يا أخي

ِسيد

ُموت

ِكس

ُبق

َبان

ُكون

َسب

ِدين

ِسين

ُروف

ِلم

ُبم

َبات

ُدوم

َفك

ِدير

َبرا

أنا قتلا   َتعالى   يا أخي

ِتين

ُروم

ِلز

ُبن

َسام

ُموت

َرن

ِجيب

ِريم

ُروز

ِمد

ُدب

َدام

ُروز

َهم

ِزيد

َبرا

أنا   قتلا   َتعالى   يا أخي

ِكيس

ُتوت

ِمس

ُهم

َفاد

ُحوش

َمس

ِزين

ِليق

ُبوك

ِتم

ُكب

َقاس

ُقوس

َون

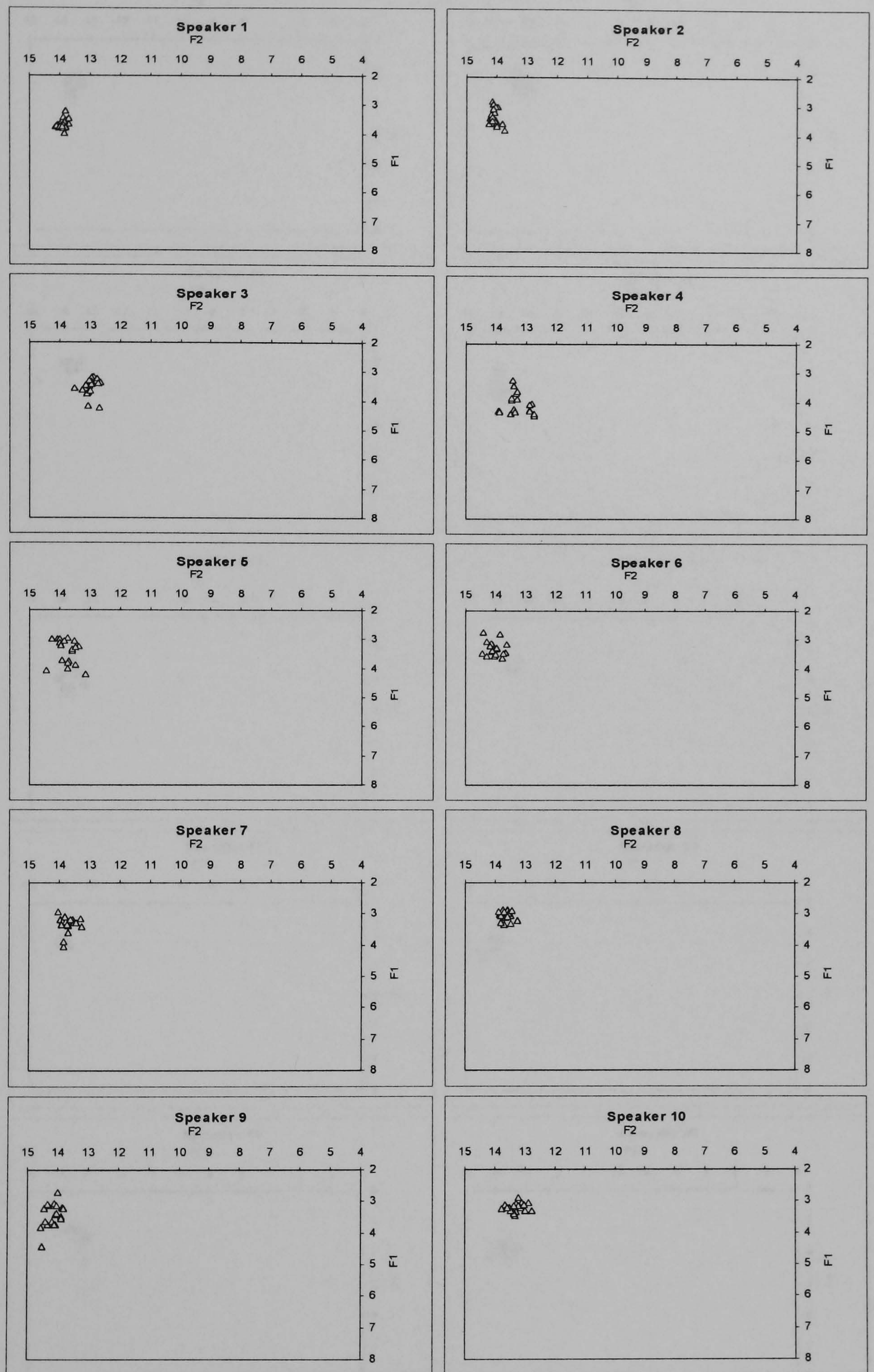
ِكيل

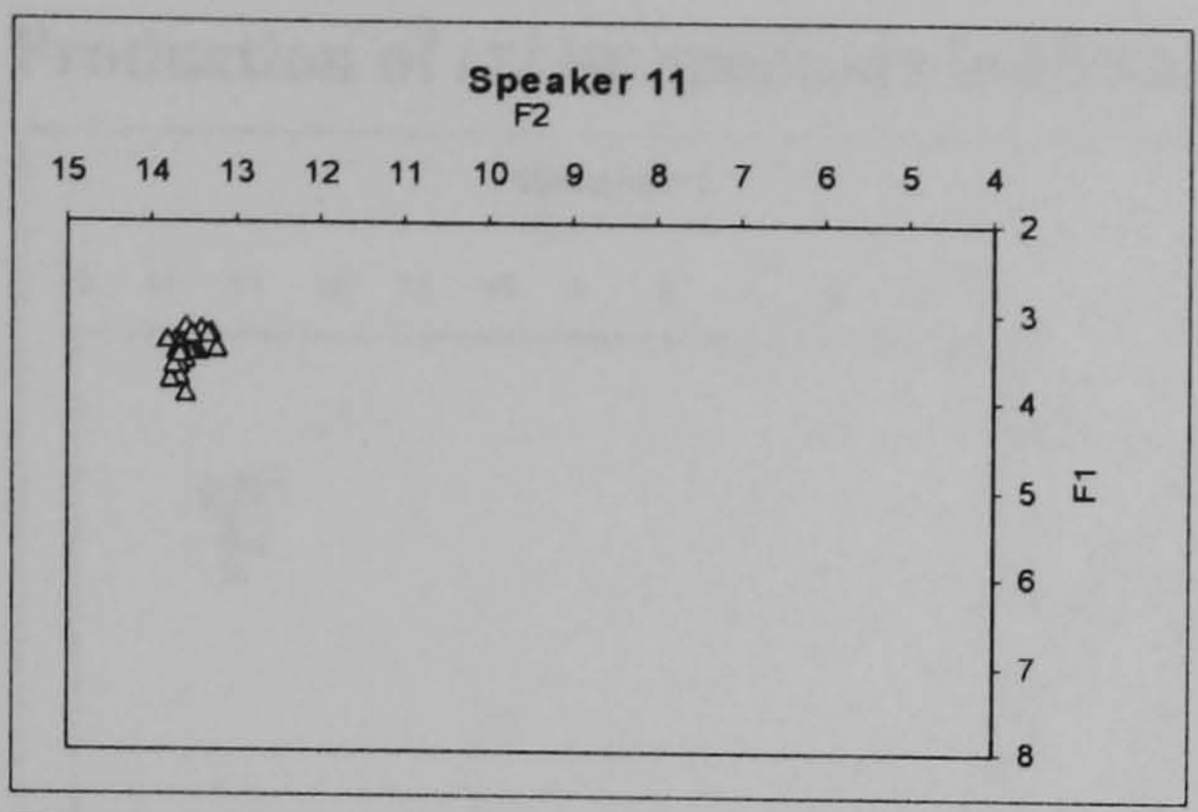
َبرا

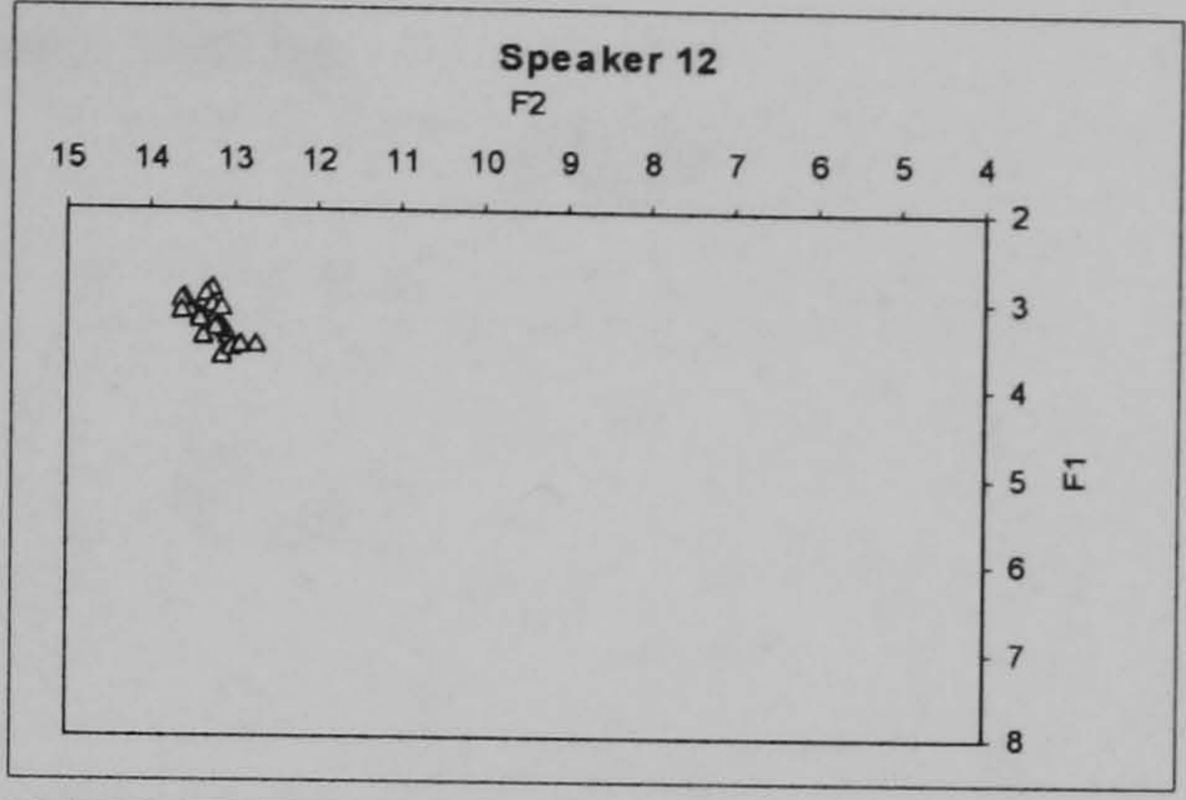
Appendix 2

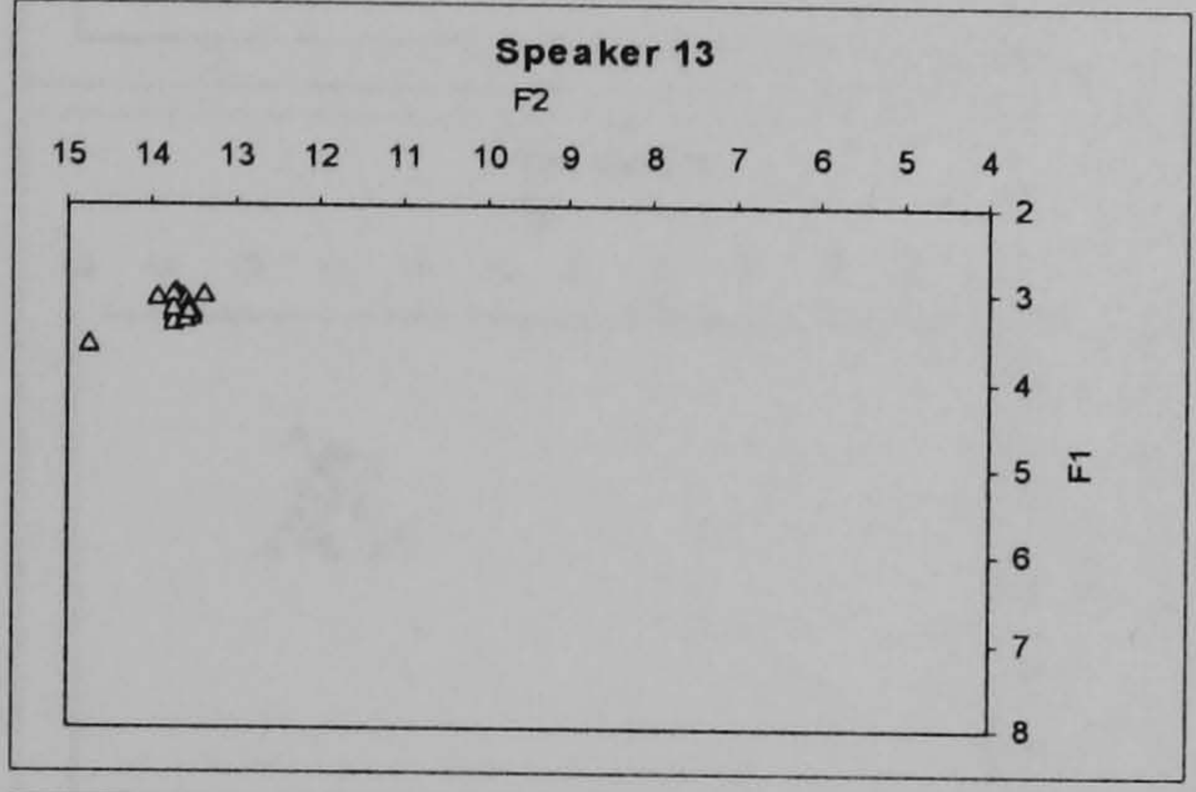
The production of individual vowels by individual speakers presented in Bark scale plane

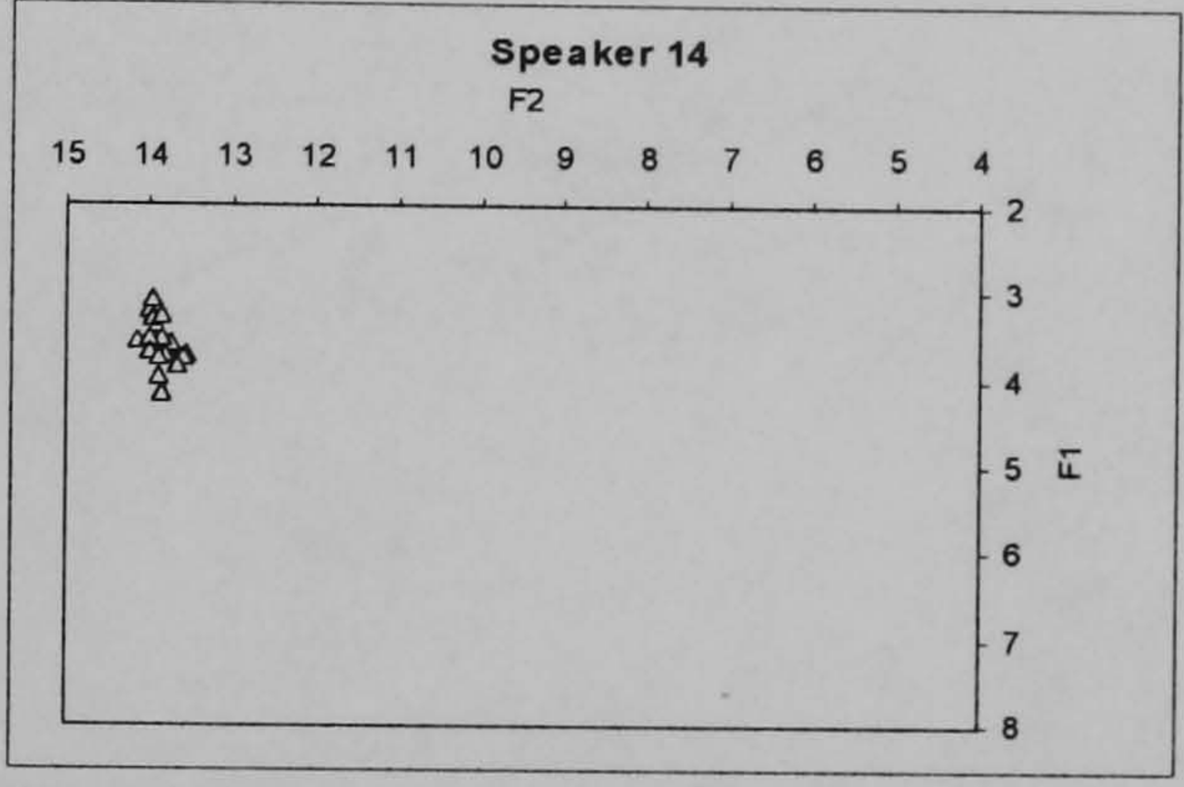
Production of /iː/ by speakers individually (Bark)

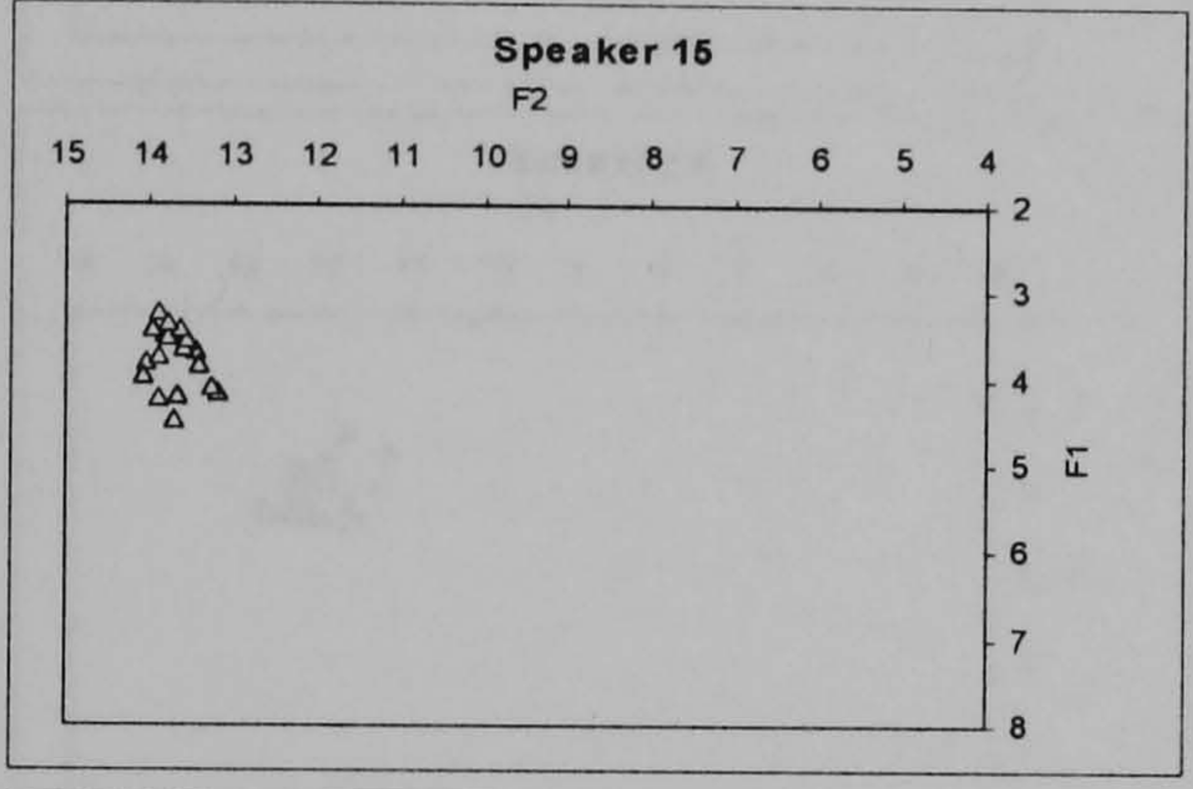


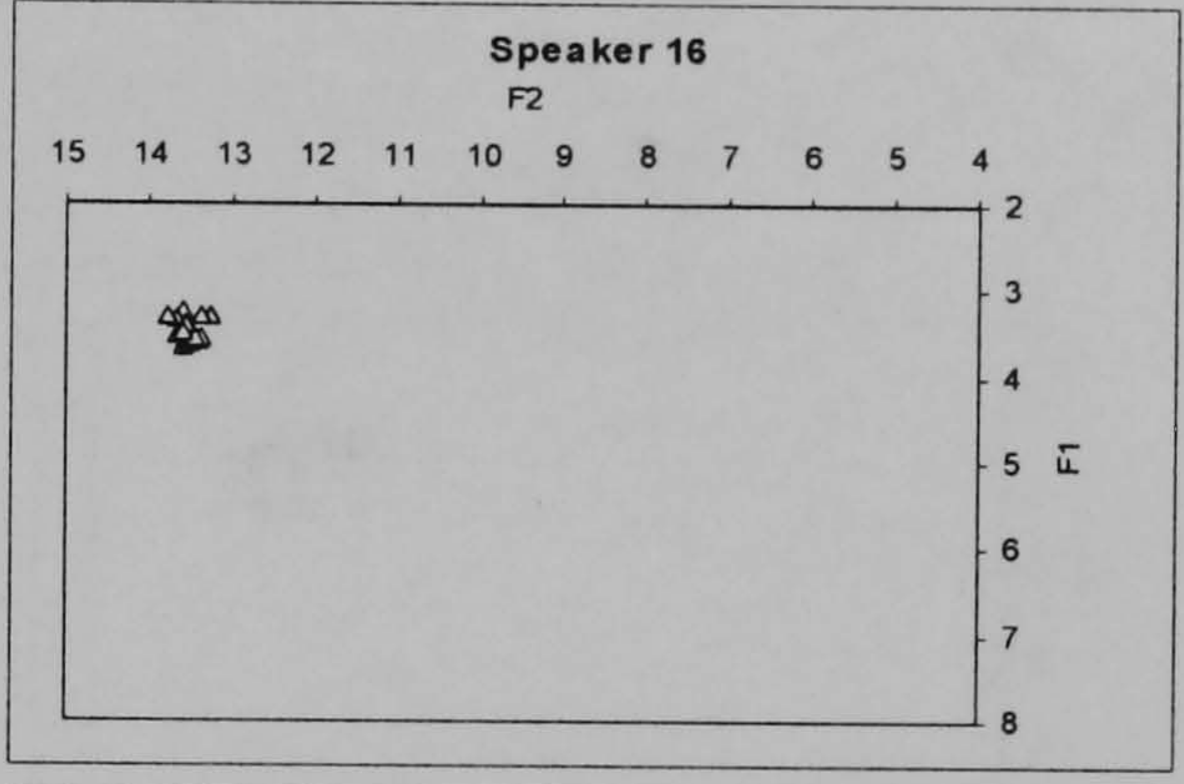

Speaker 11
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

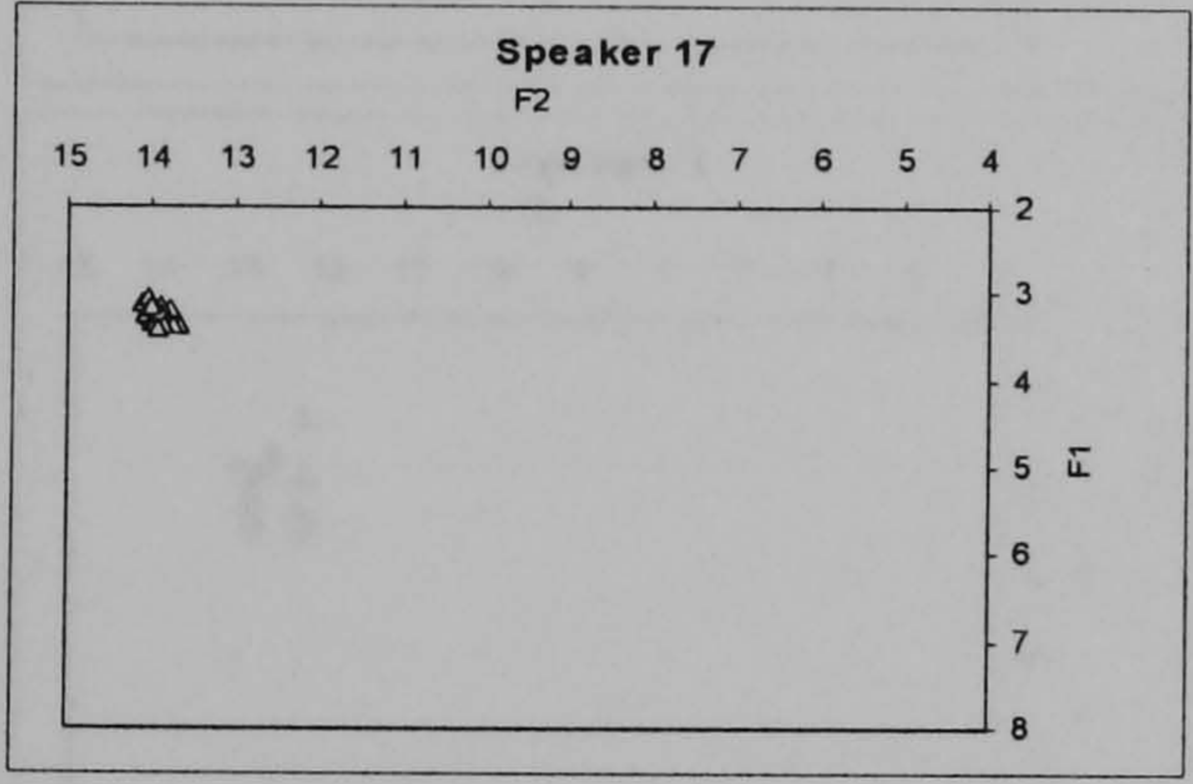

Speaker 12
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

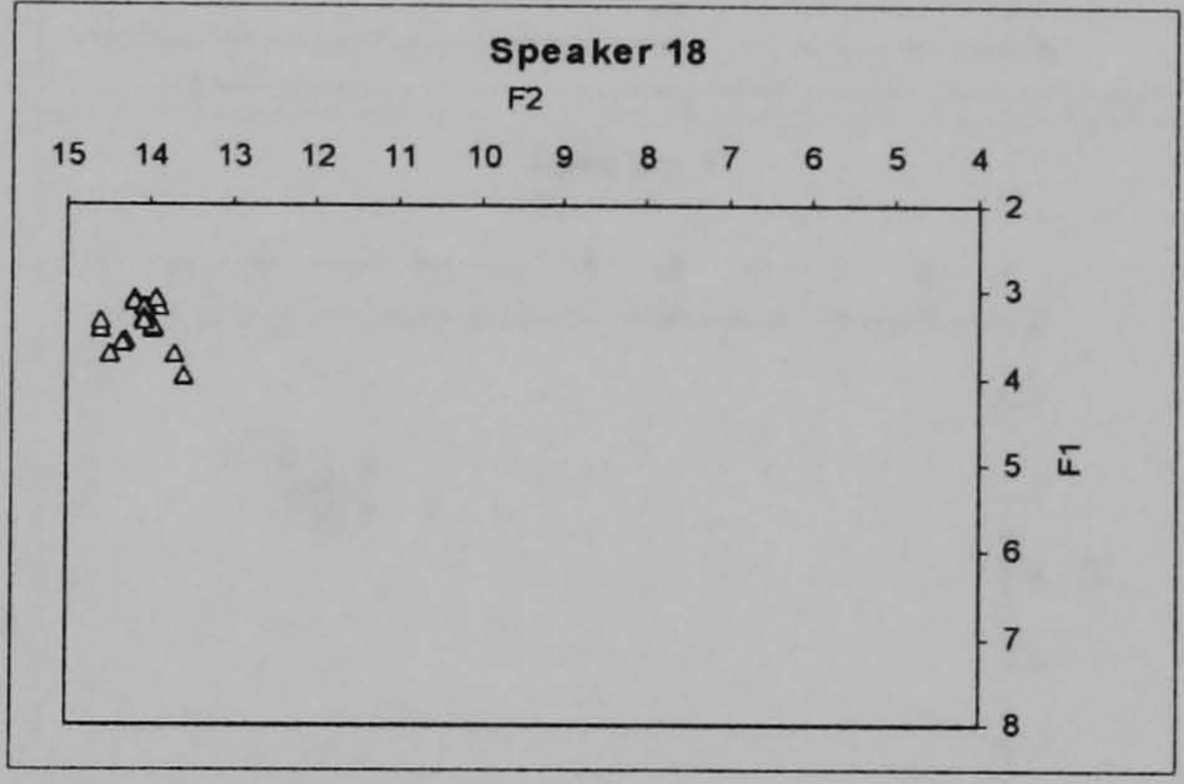

Speaker 13
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

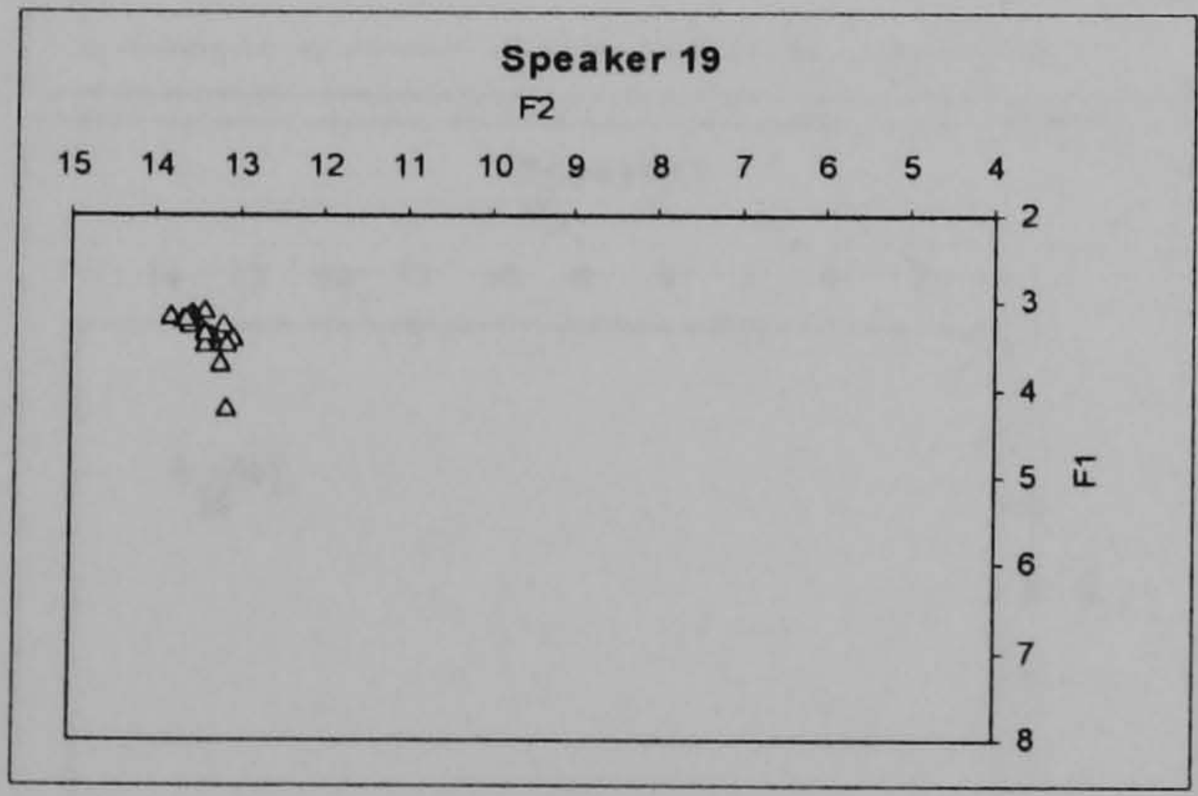

Speaker 14
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

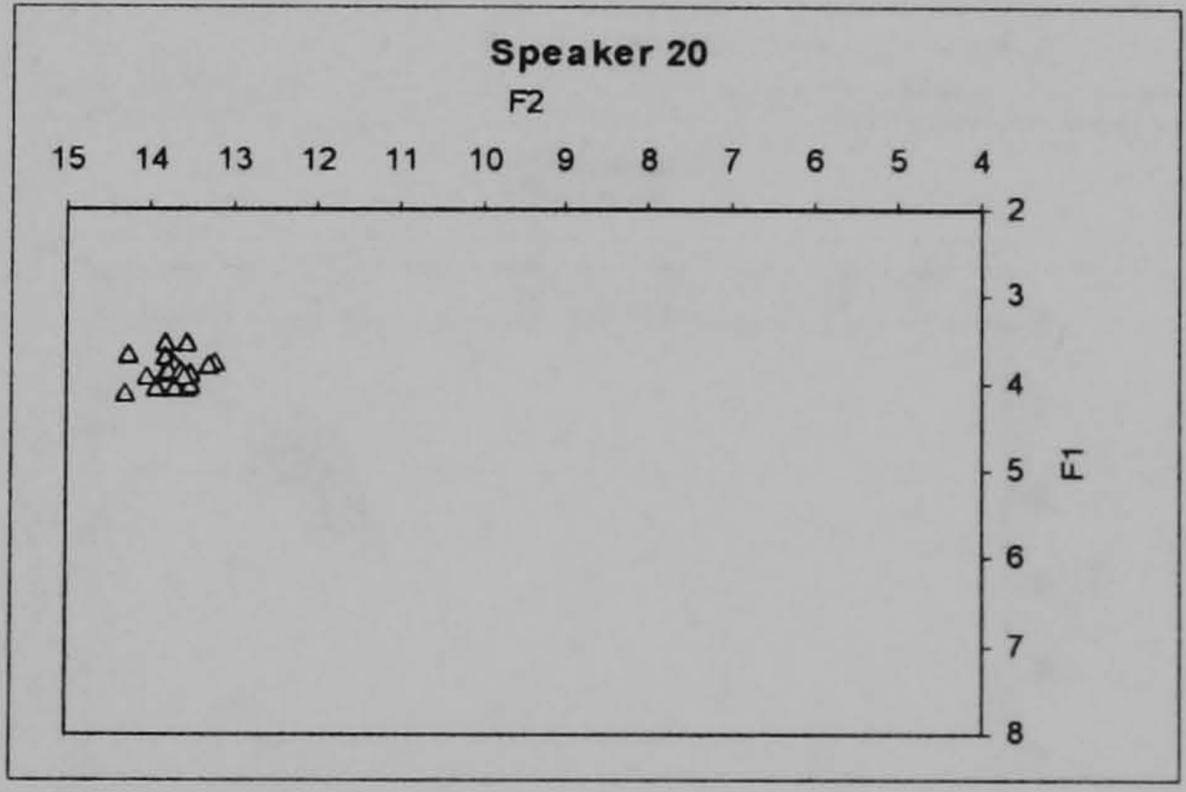

Speaker 15
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1


Speaker 16
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

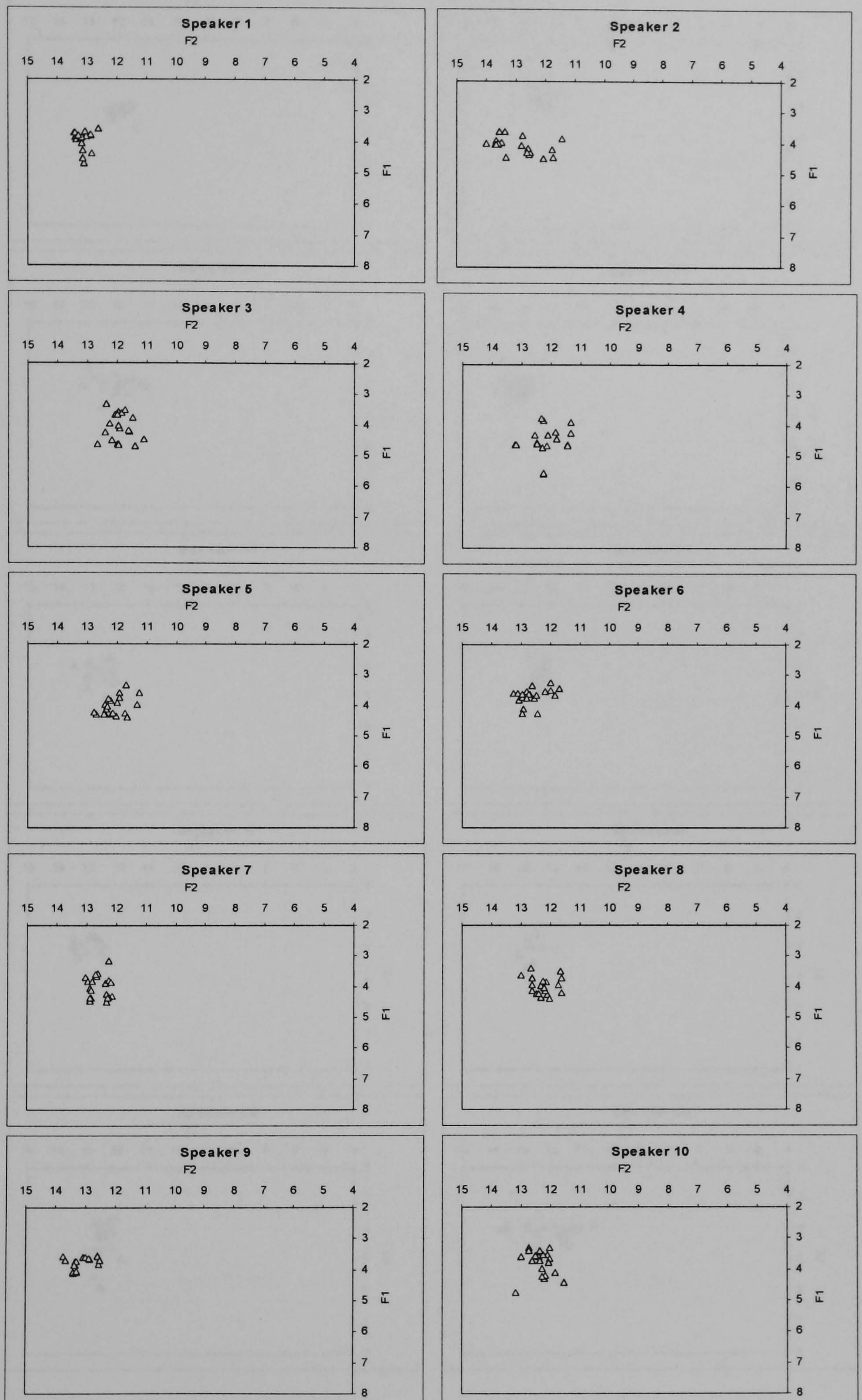

Speaker 17
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

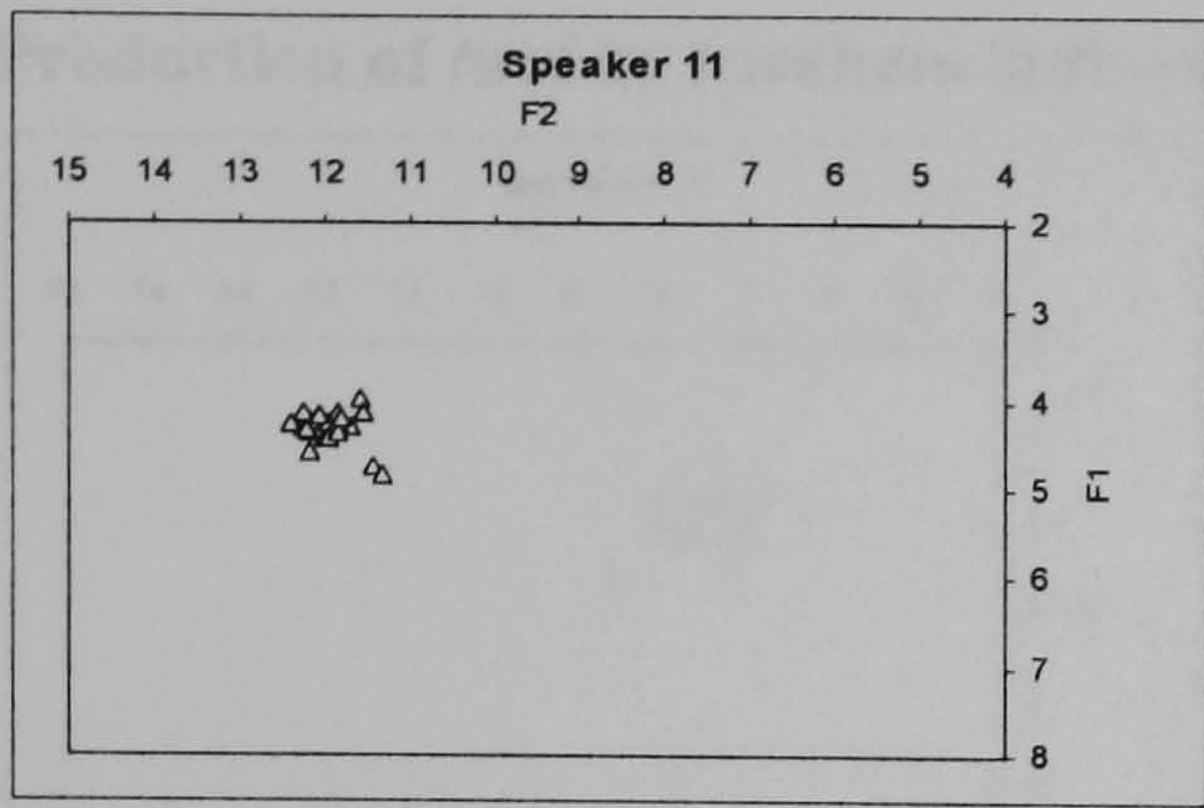

Speaker 18
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

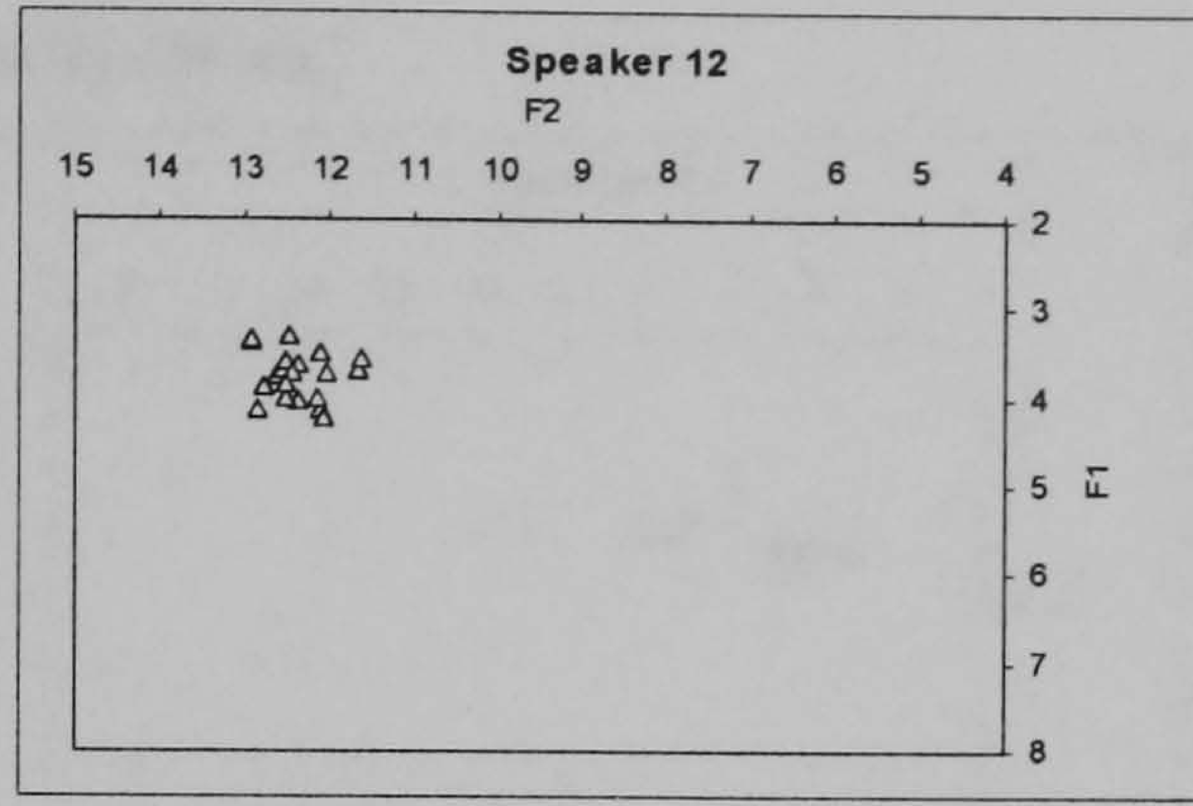

Speaker 19
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

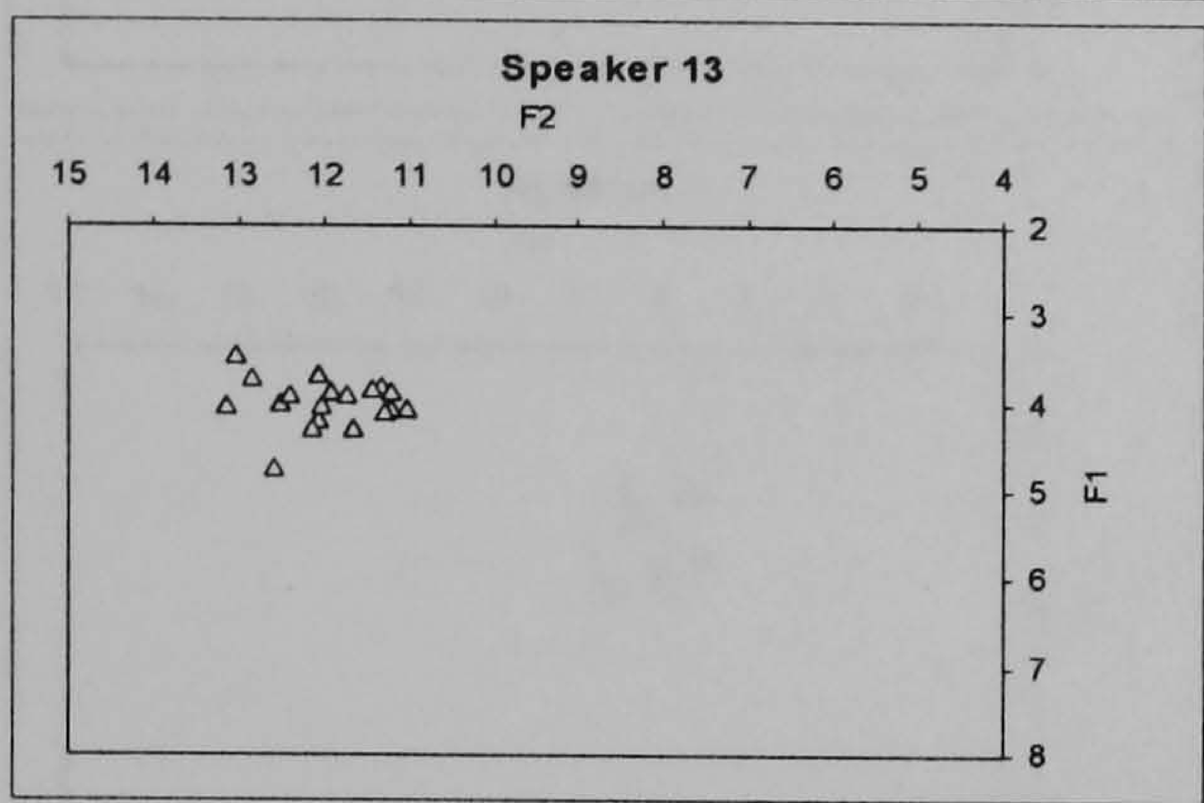

Speaker 20
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

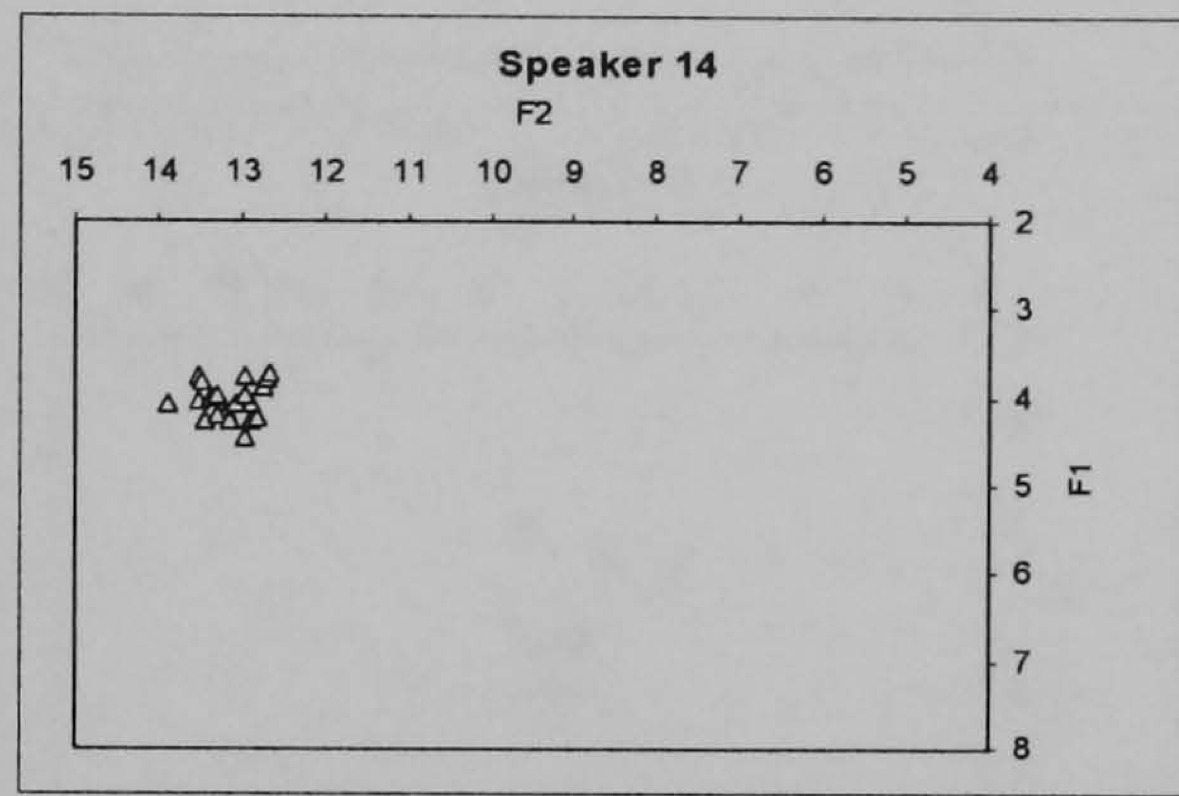
Production of /ɪ/ by speakers individually (Bark)

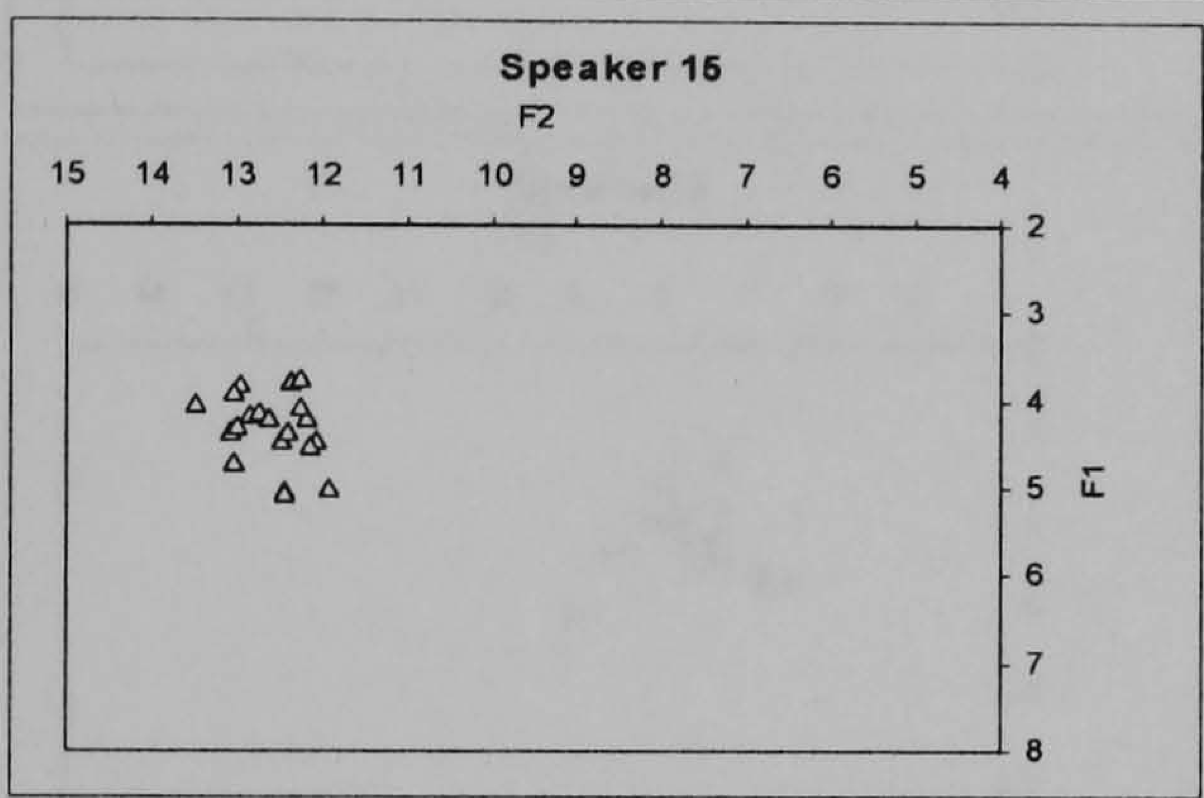


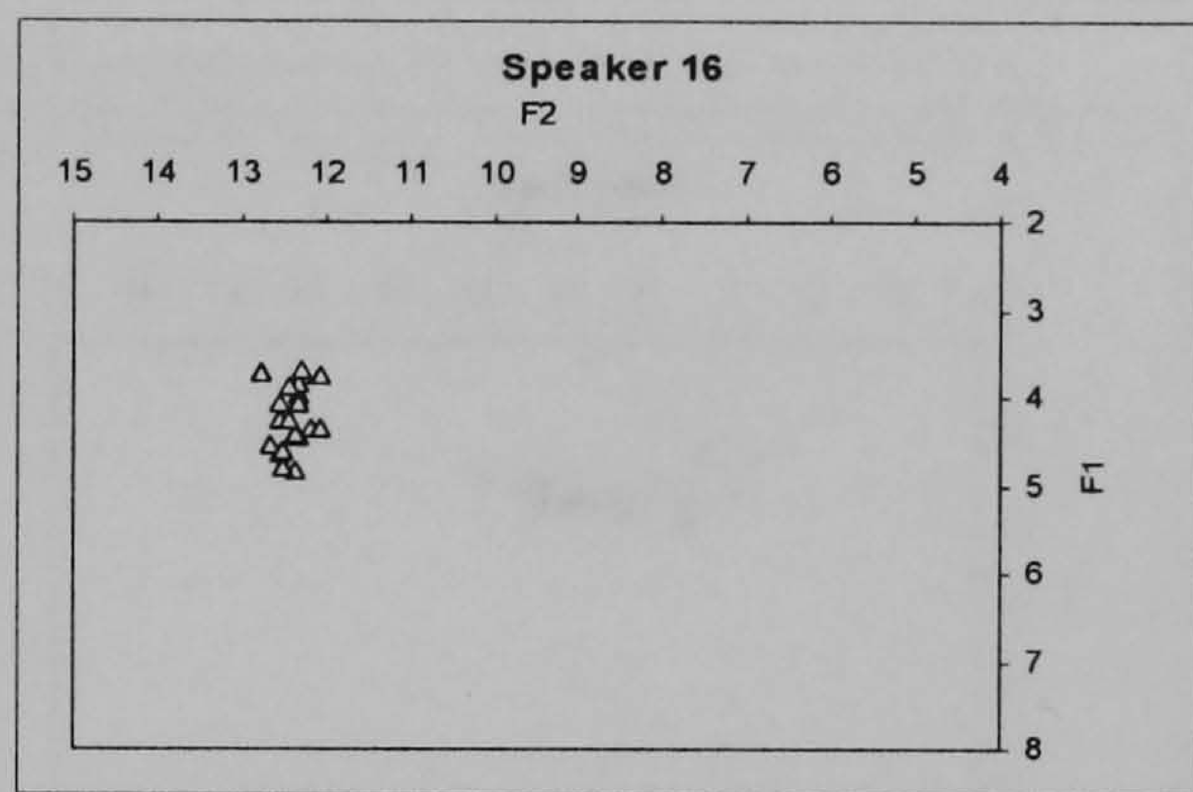

Speaker 11
F2
F1

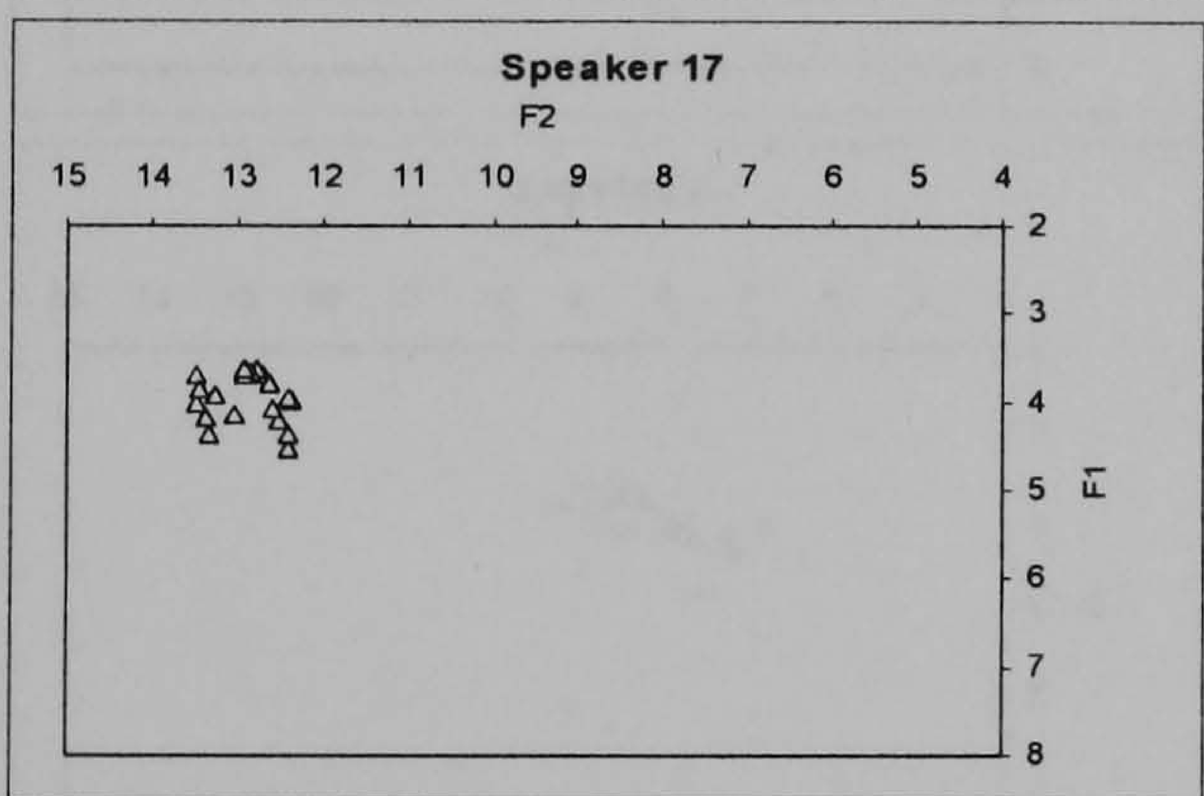

Speaker 12
F2
F1

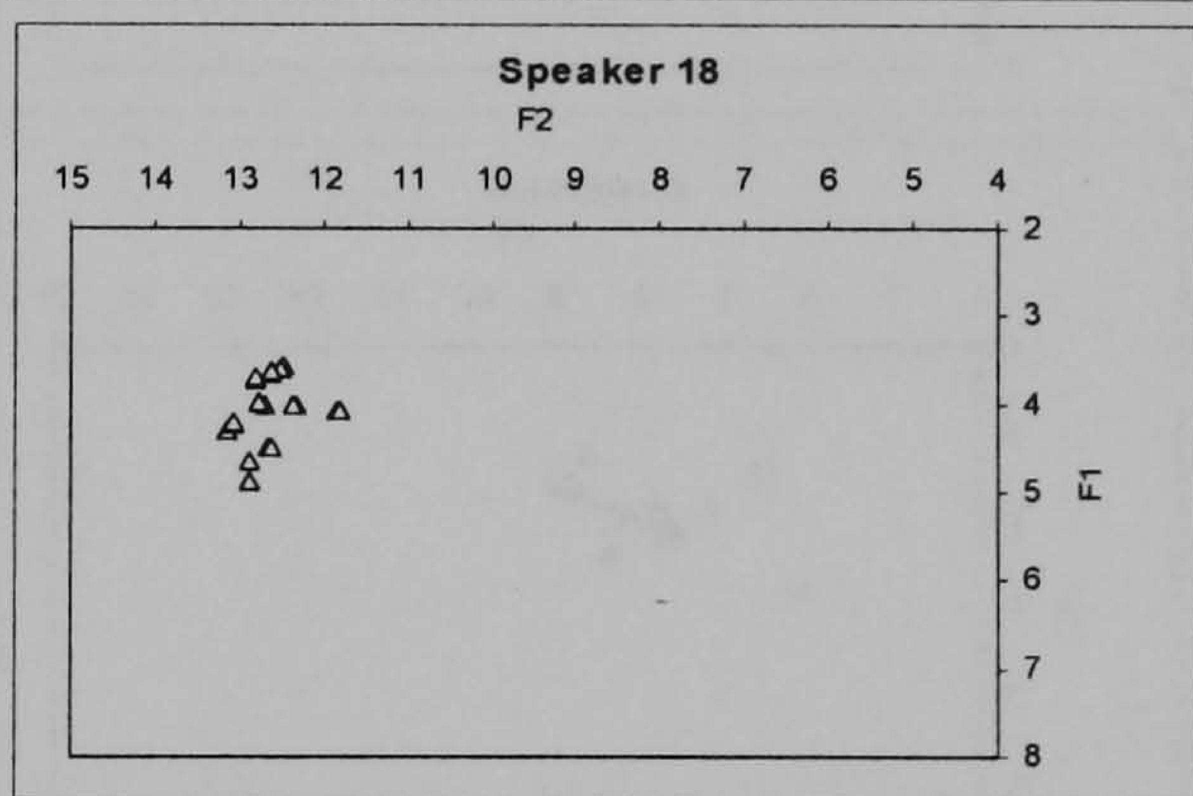

Speaker 13
F2
F1

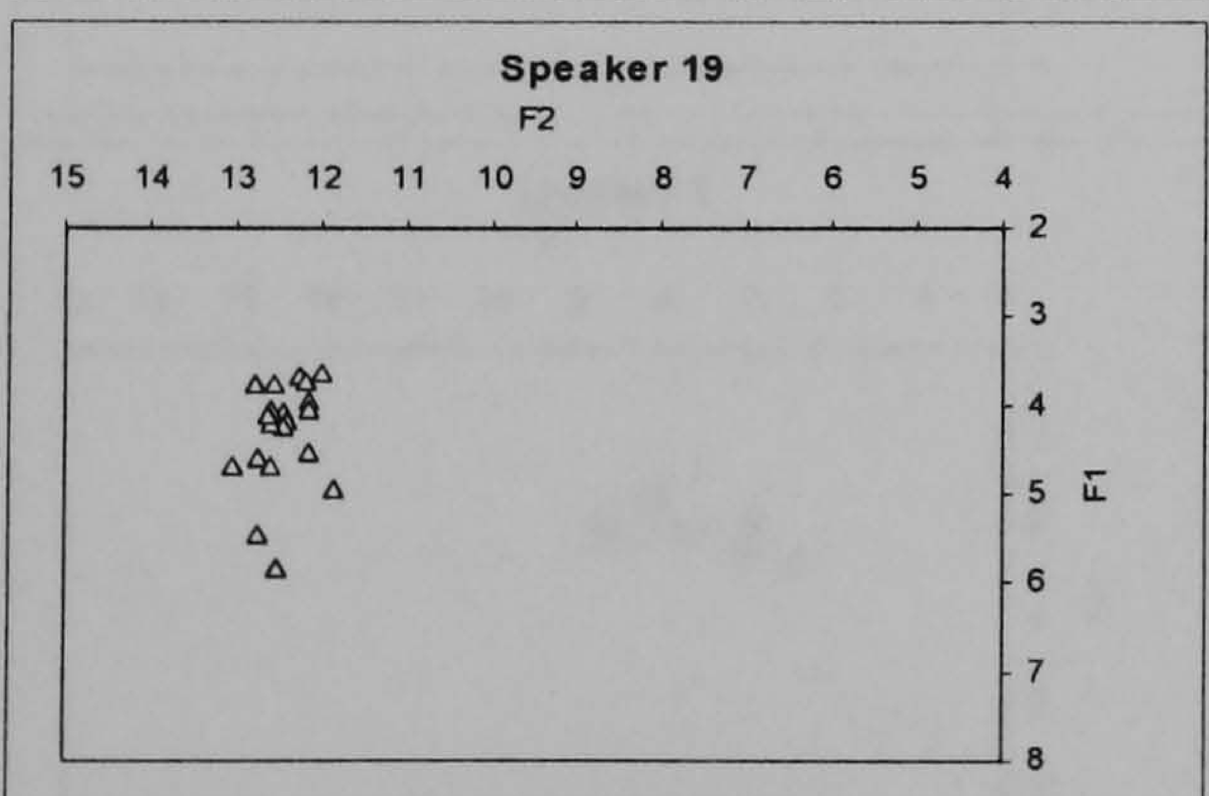

Speaker 14
F2
F1

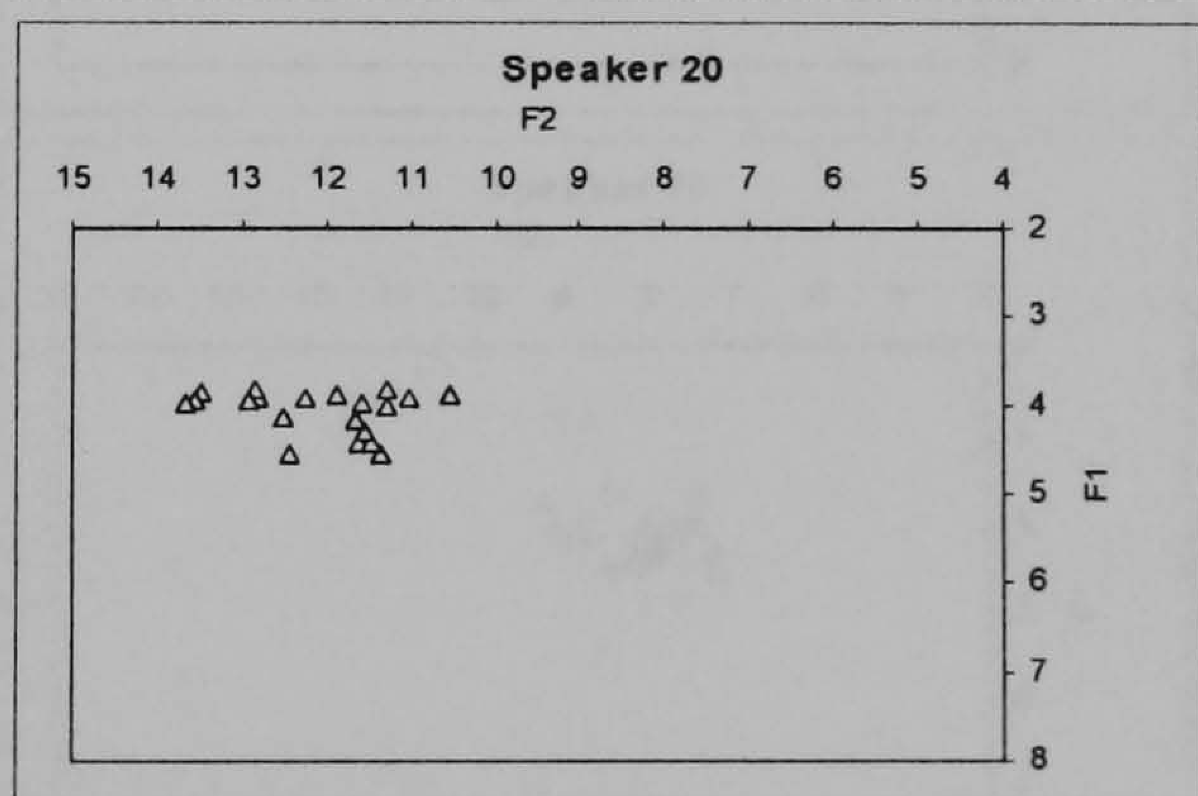

Speaker 15
F2
F1


Speaker 16
F2
F1

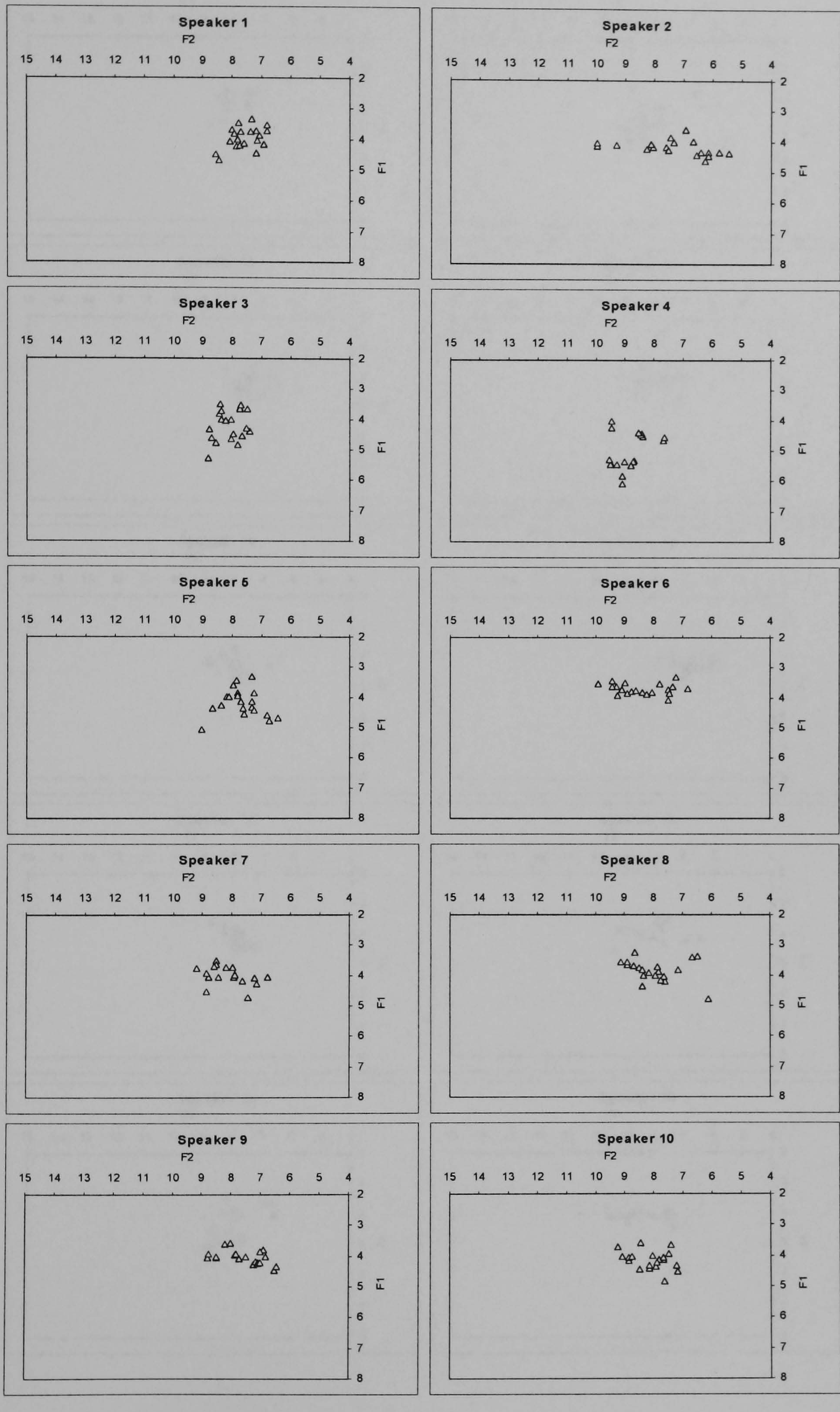

Speaker 17
F2
F1

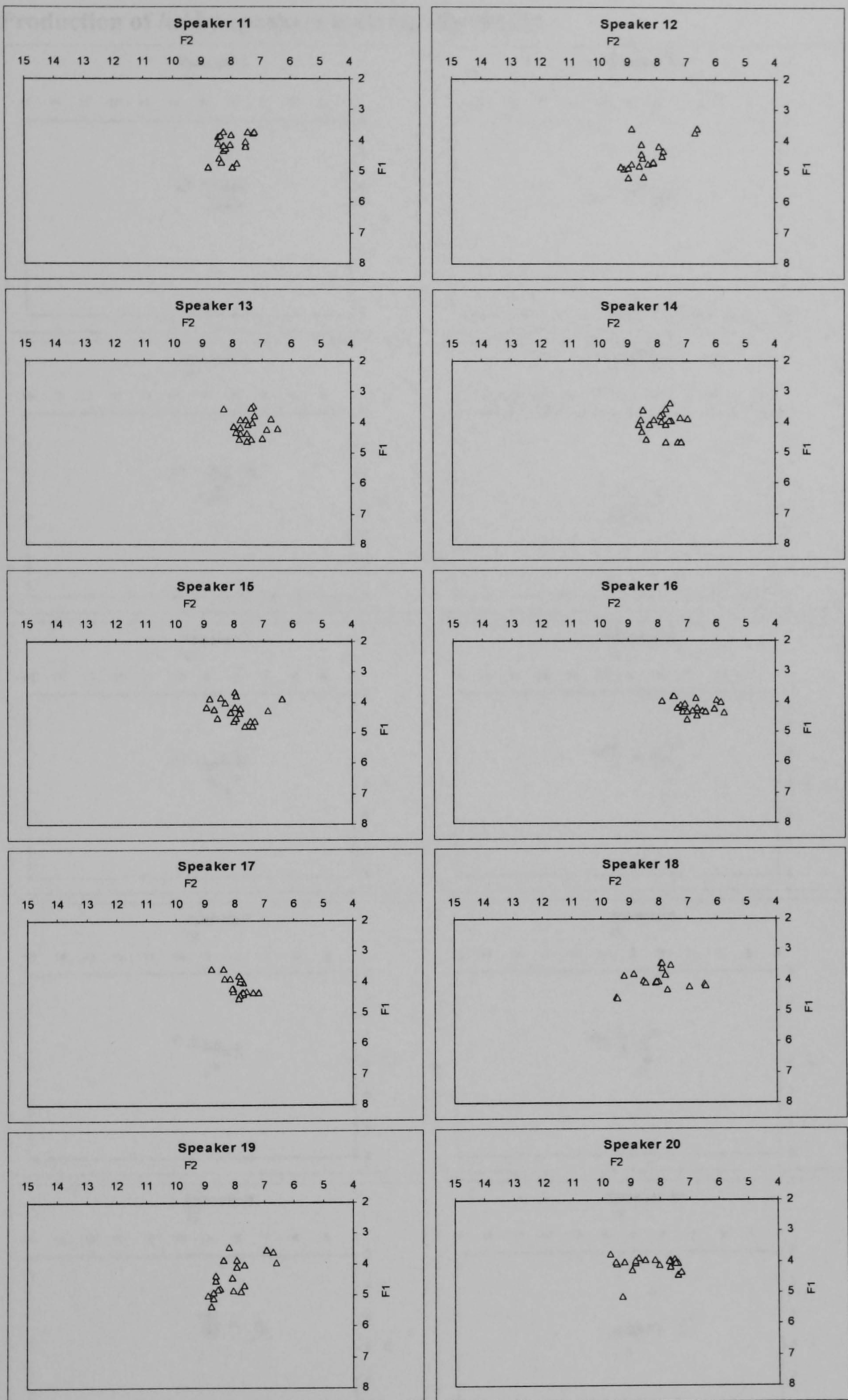

Speaker 18
F2
F1


Speaker 19
F2
F1

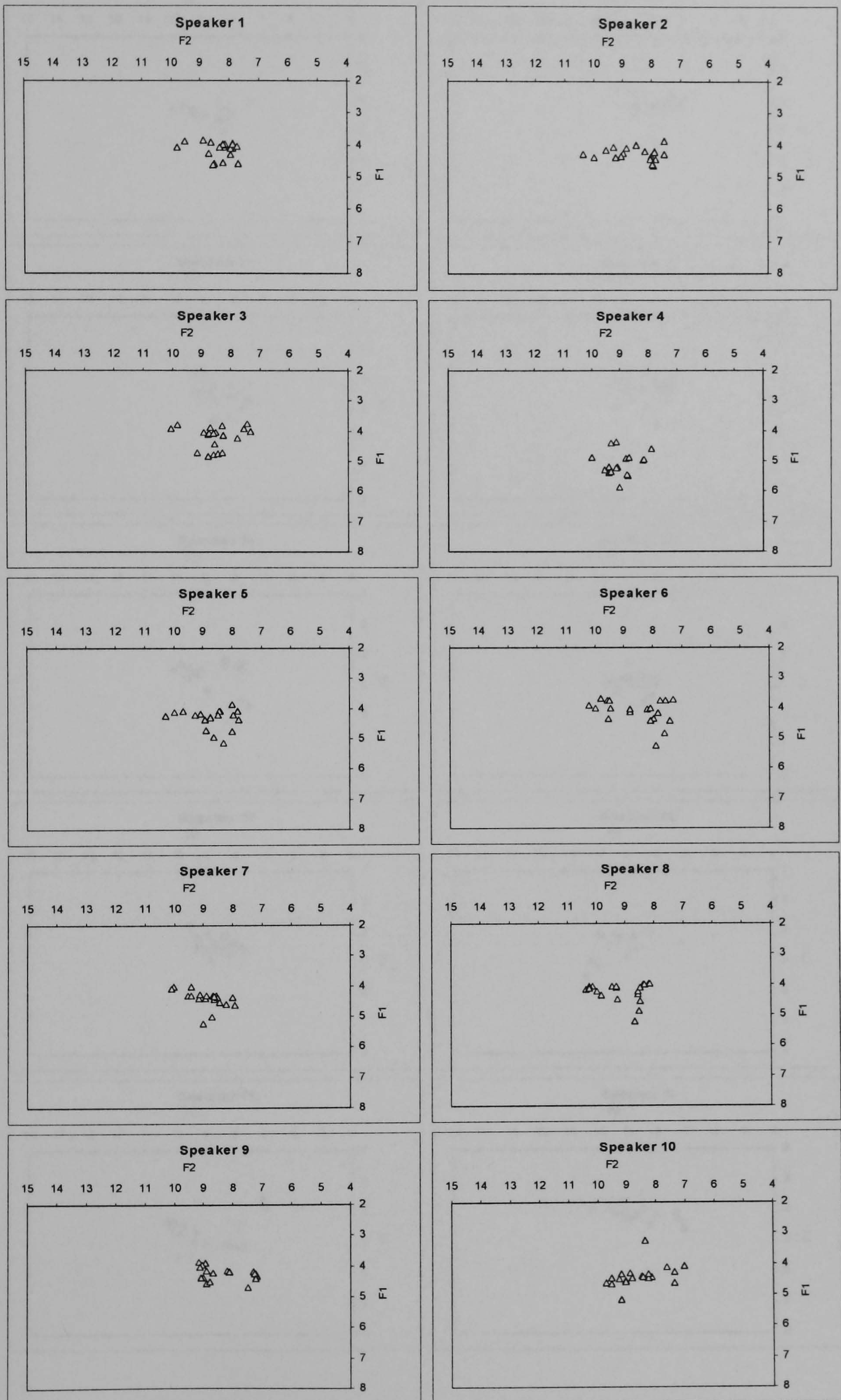

Speaker 20
F2
F1

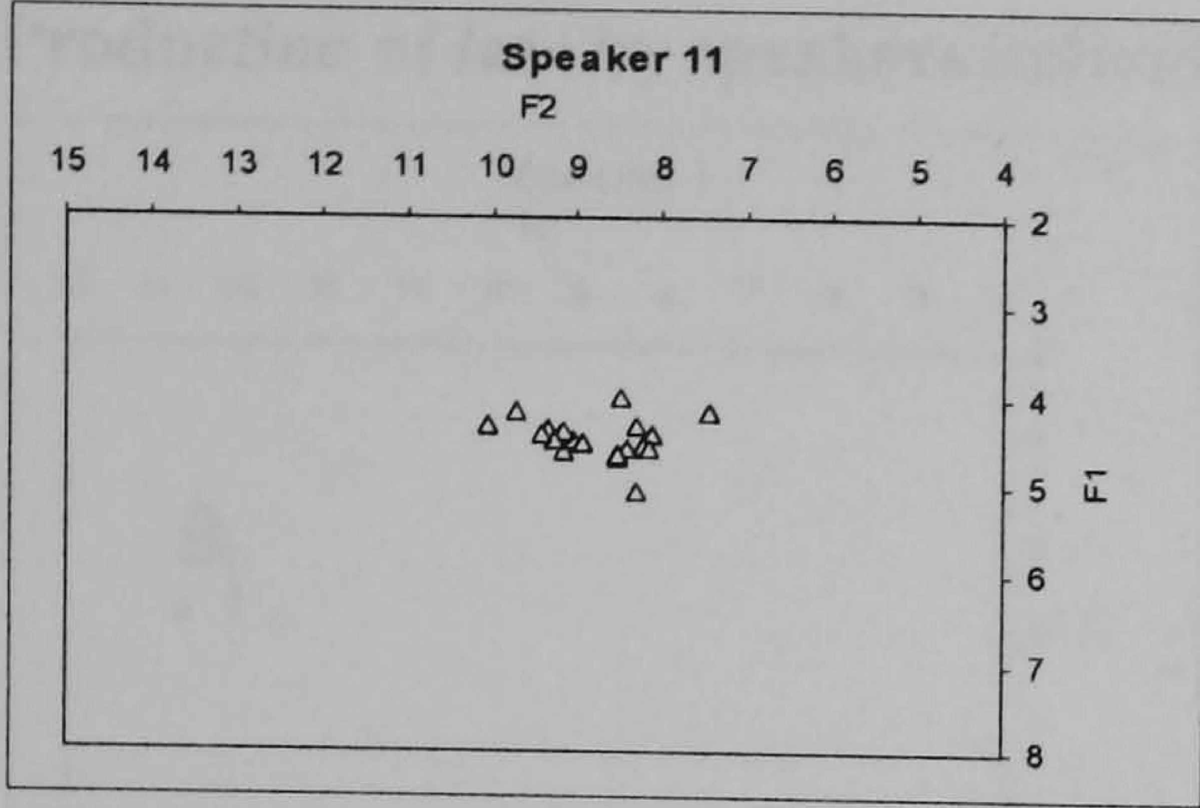
Production of /u:/ by speakers individually (Bark)

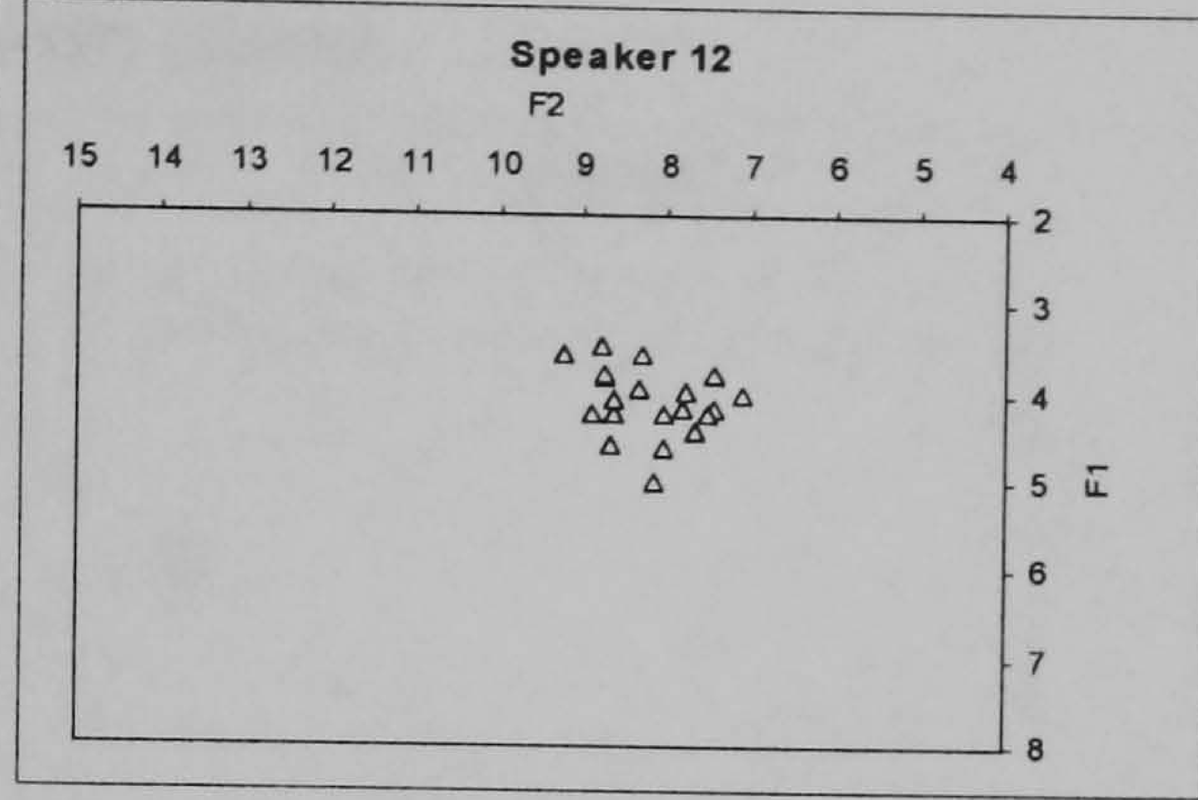


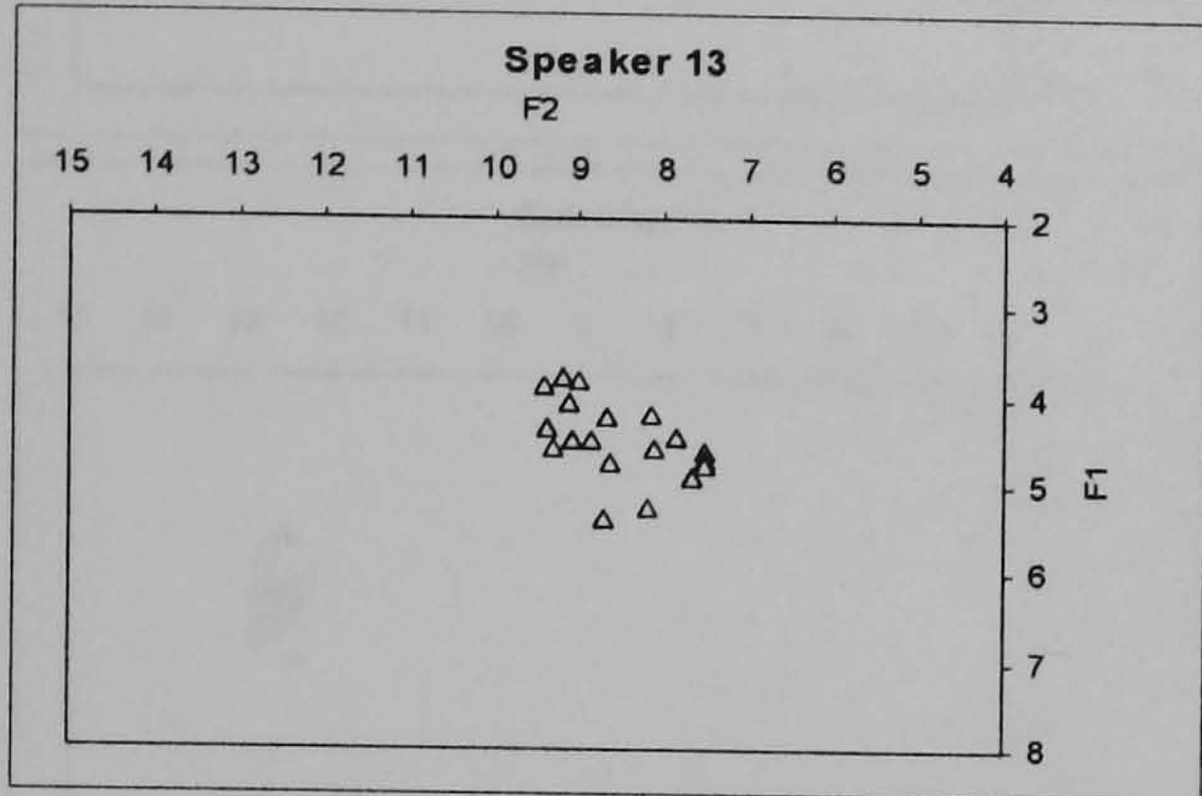

Speaker 11
Speaker 12
Speaker 13
Speaker 14
Speaker 15
Speaker 16
Speaker 17
Speaker 18
Speaker 19
Speaker 20
F2
F1
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8

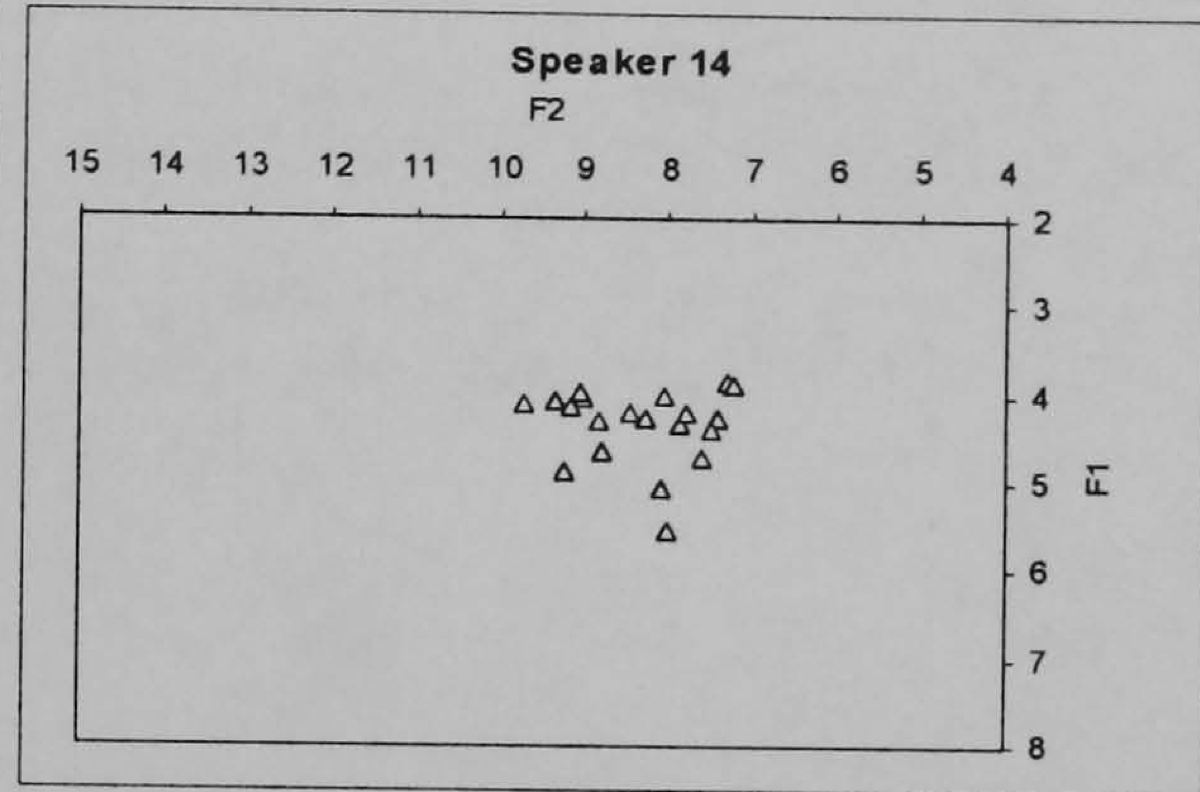
Production of /ʊ/ by speakers individually (Bark)

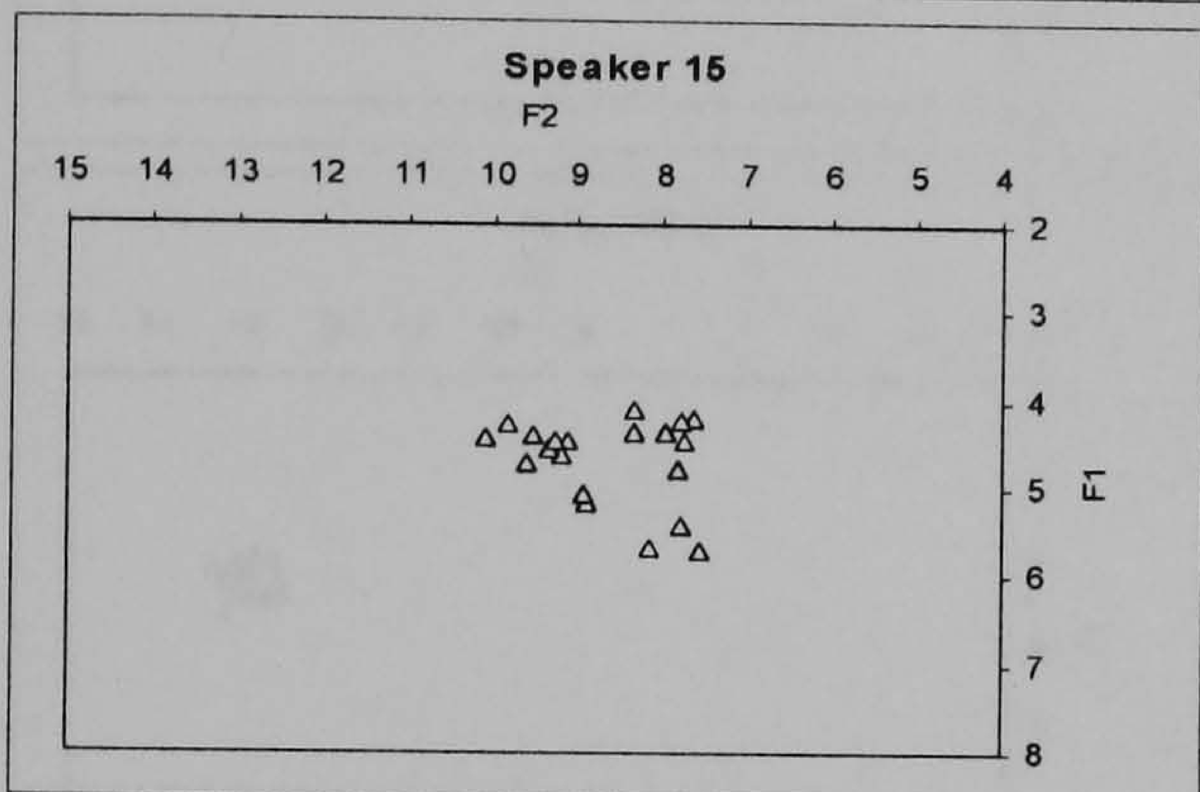


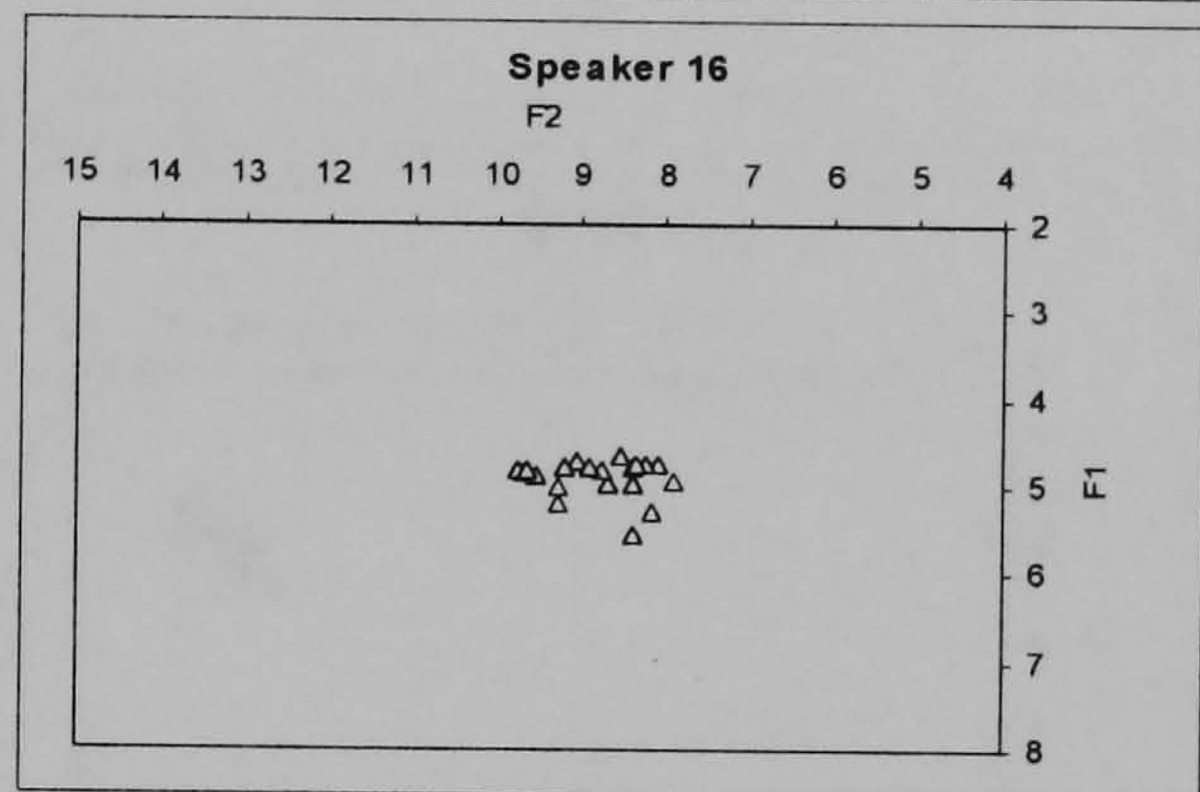

Speaker 11
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

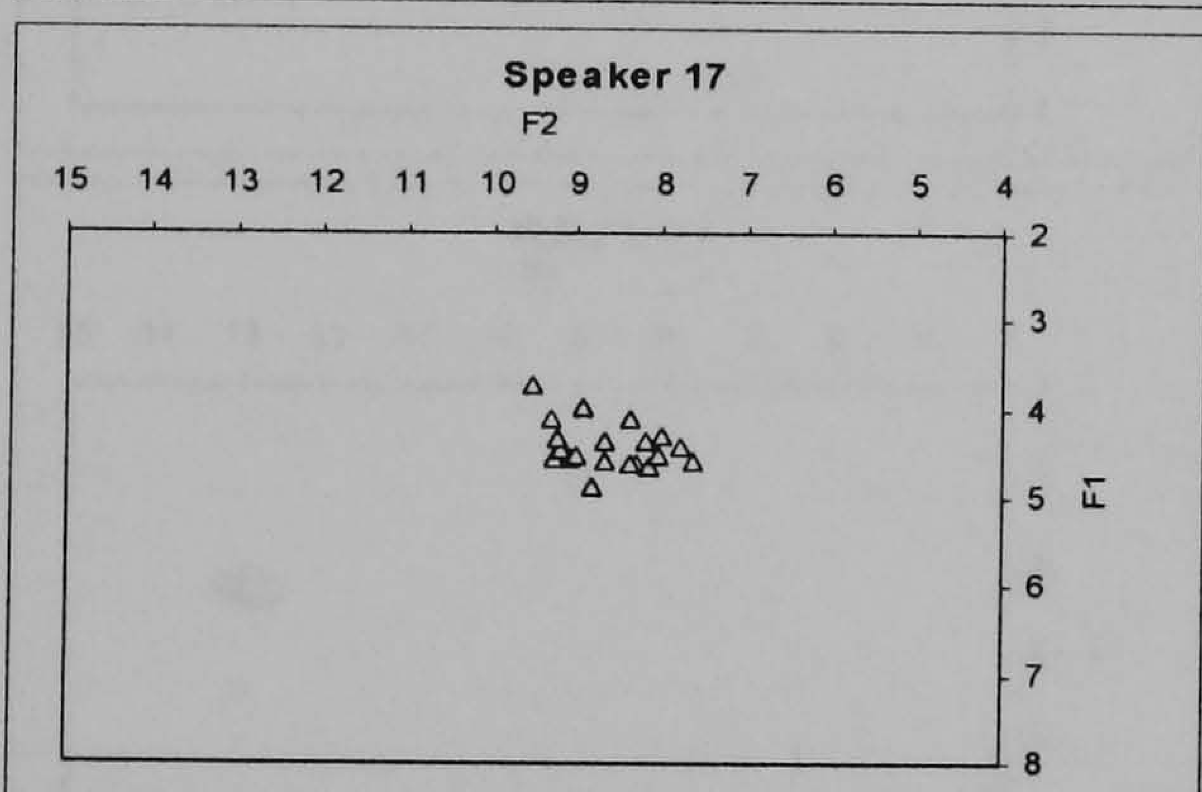

Speaker 12
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

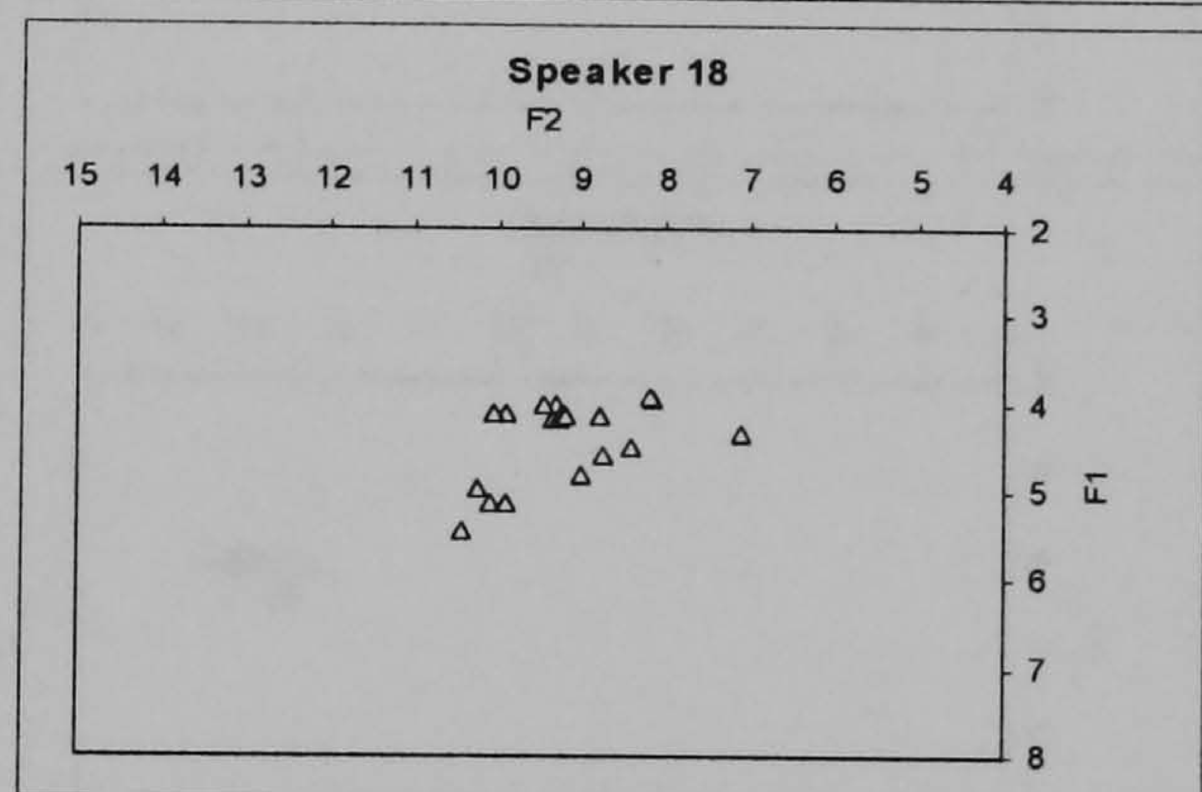

Speaker 13
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

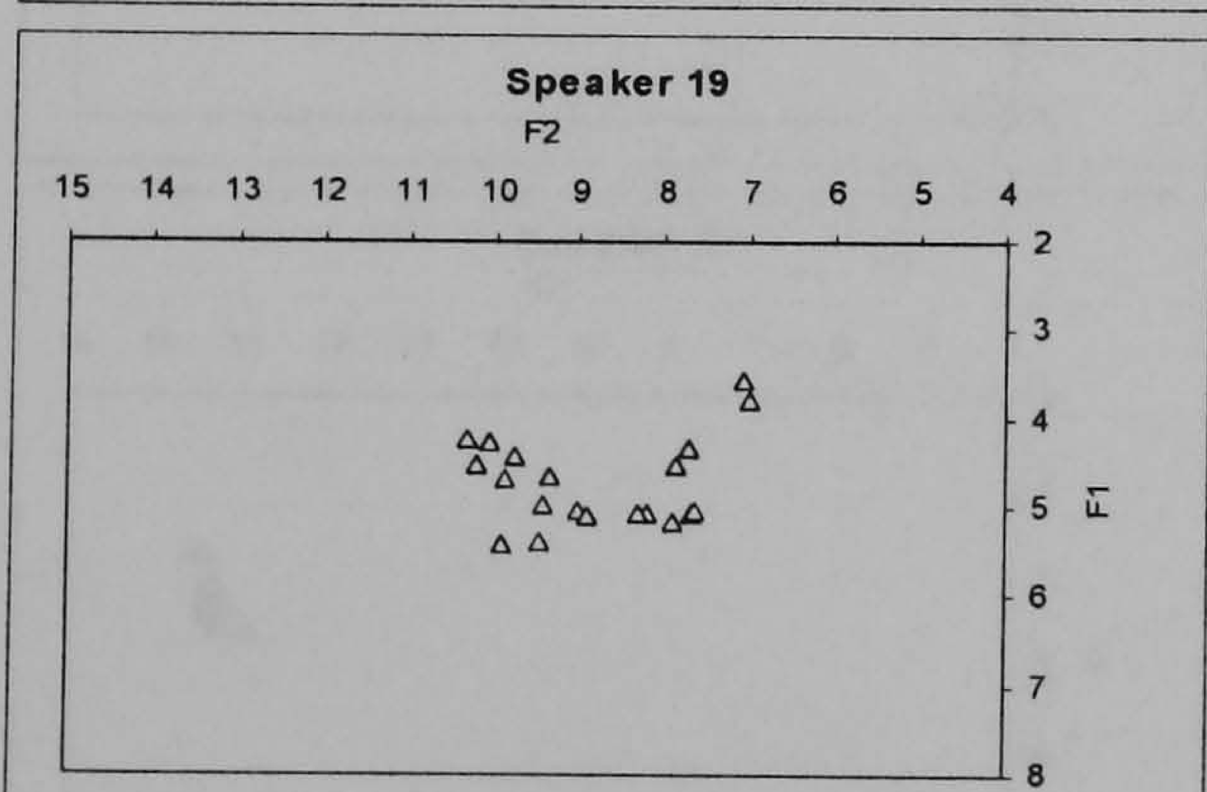

Speaker 14
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

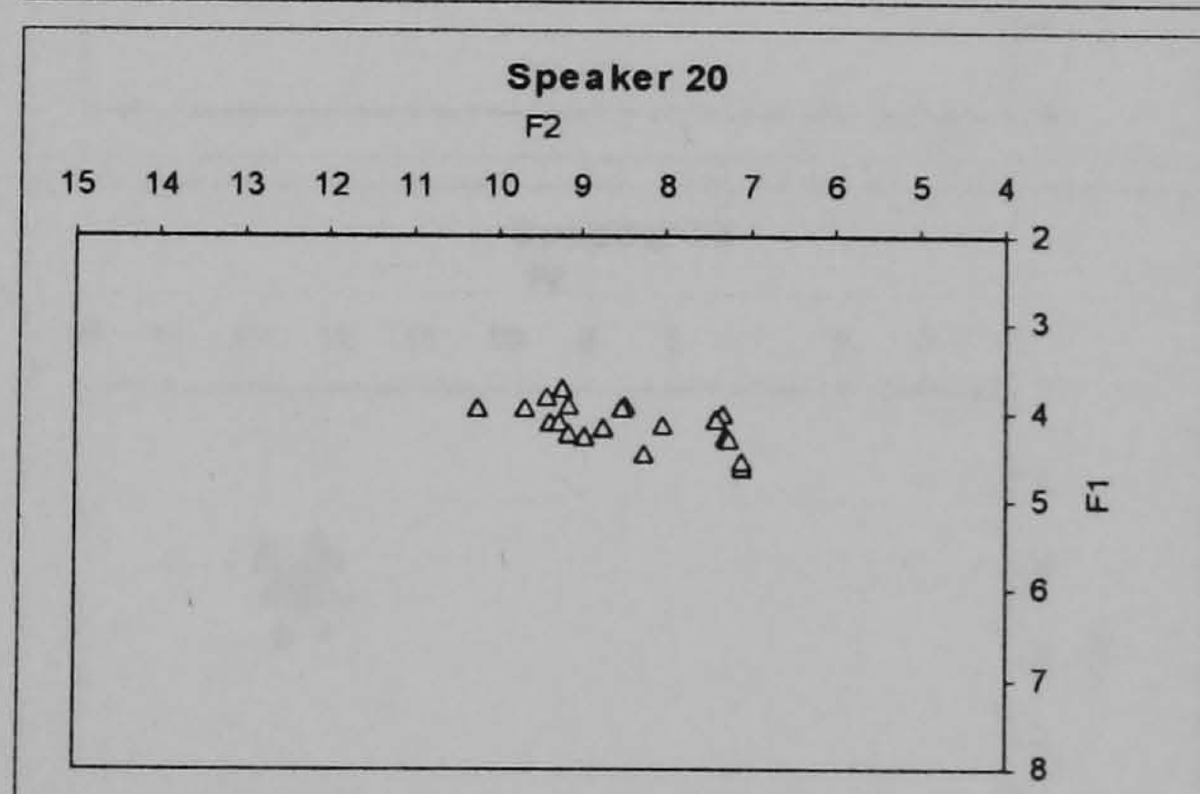

Speaker 15
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1


Speaker 16
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

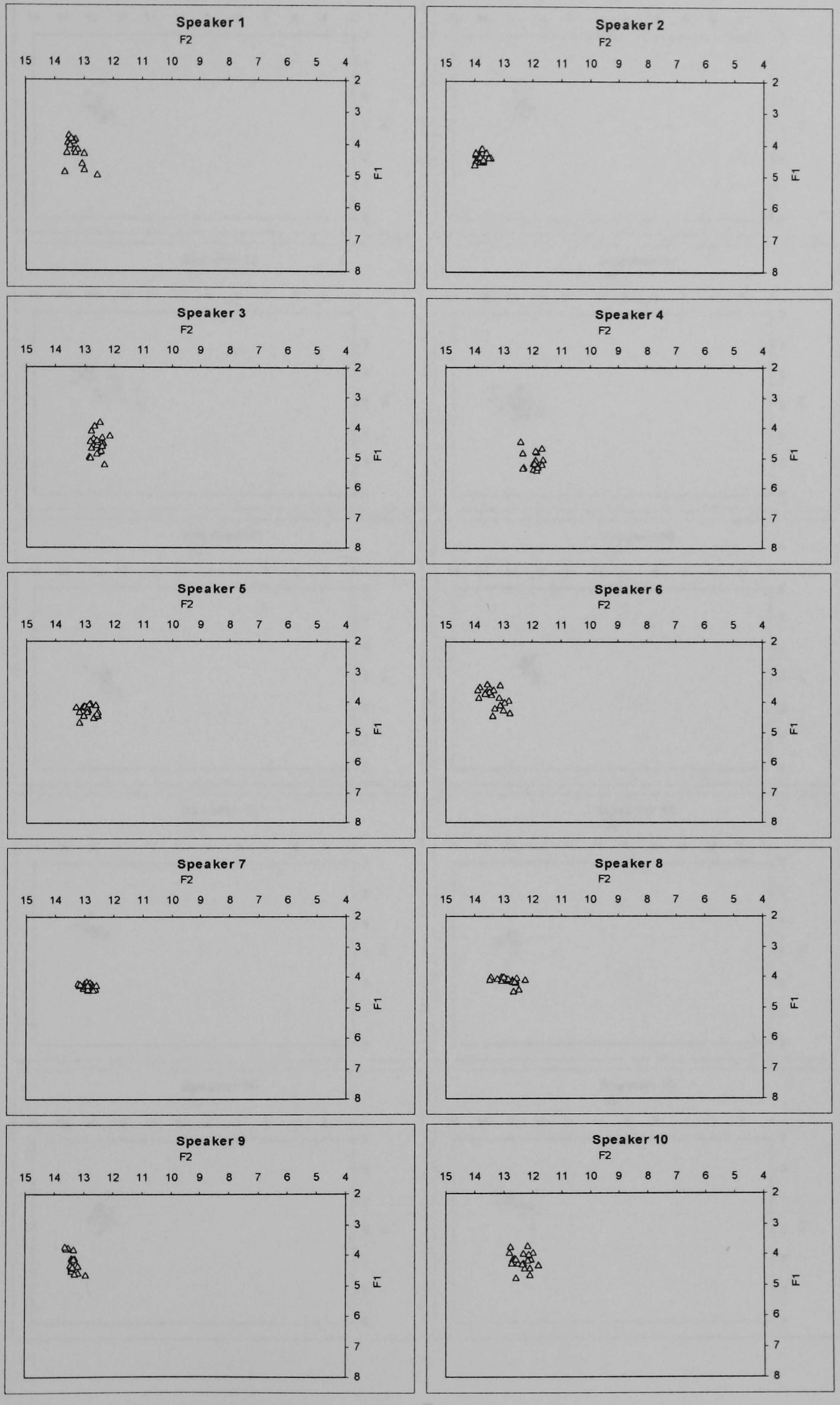

Speaker 17
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

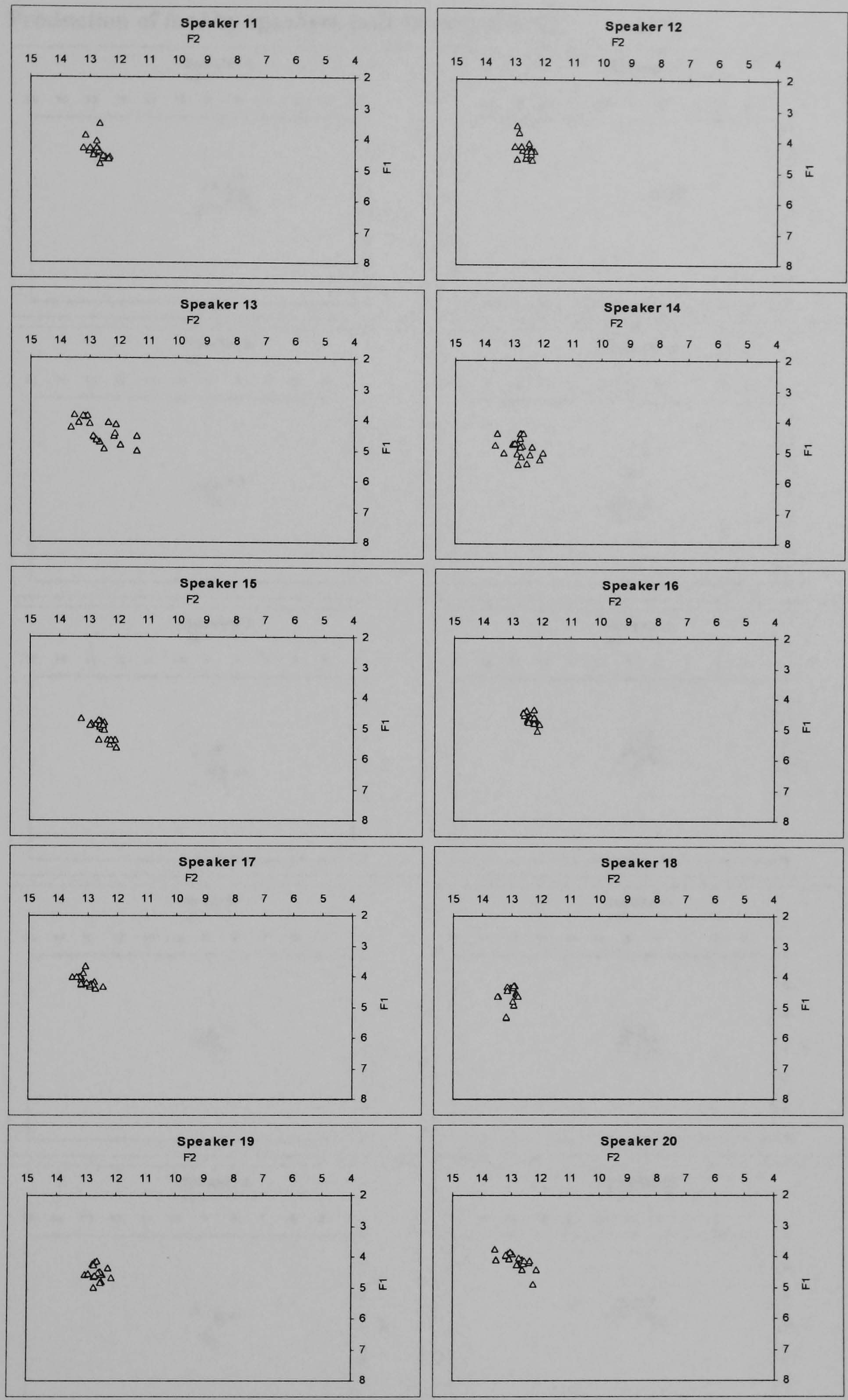

Speaker 18
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1


Speaker 19
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

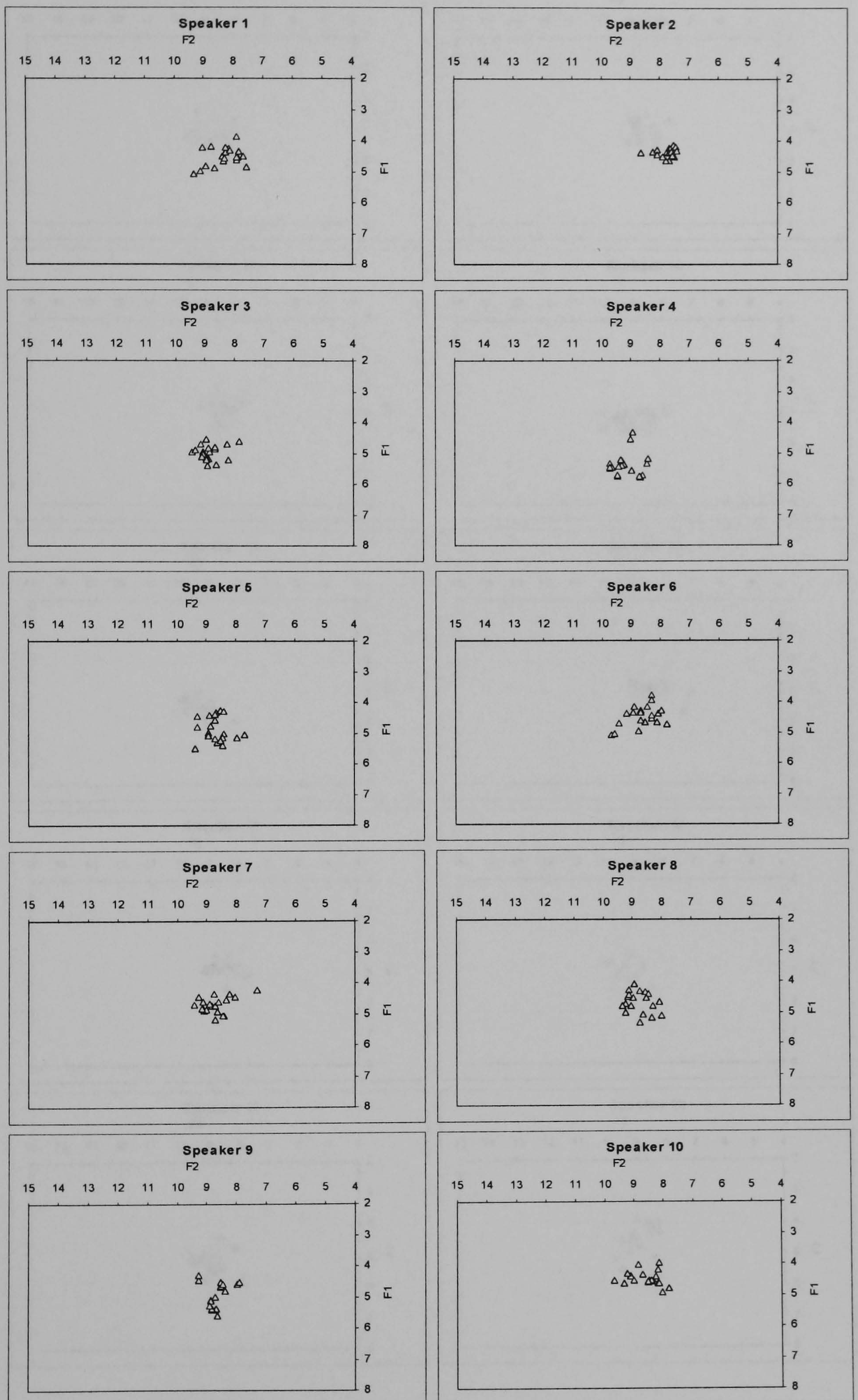

Speaker 20
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

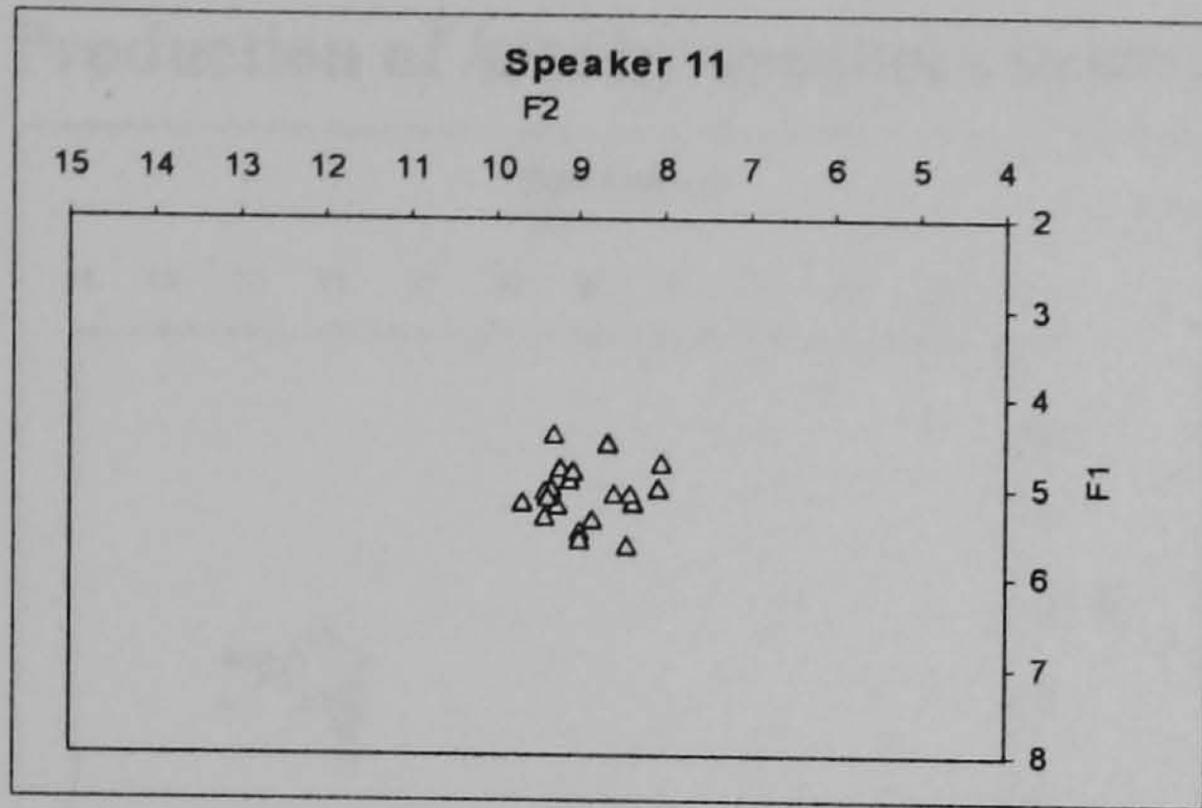
Production of /e:/ by speakers individually (Bark)

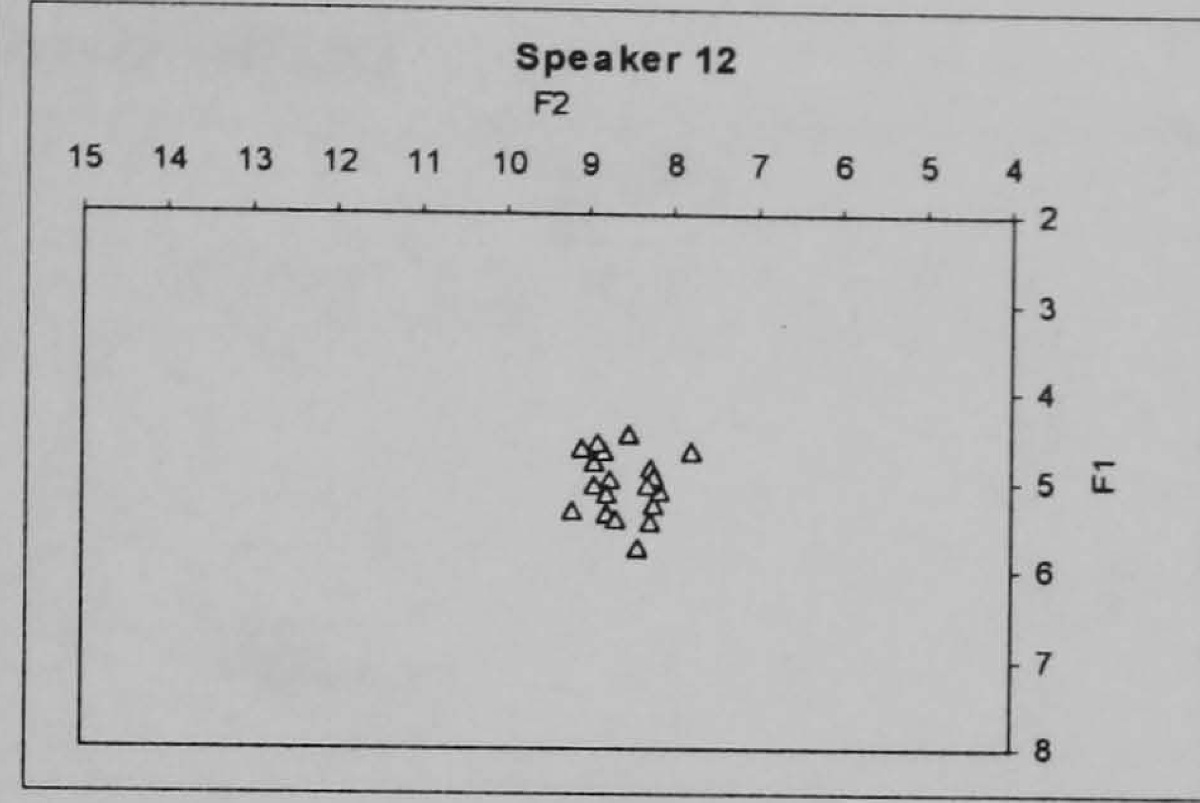


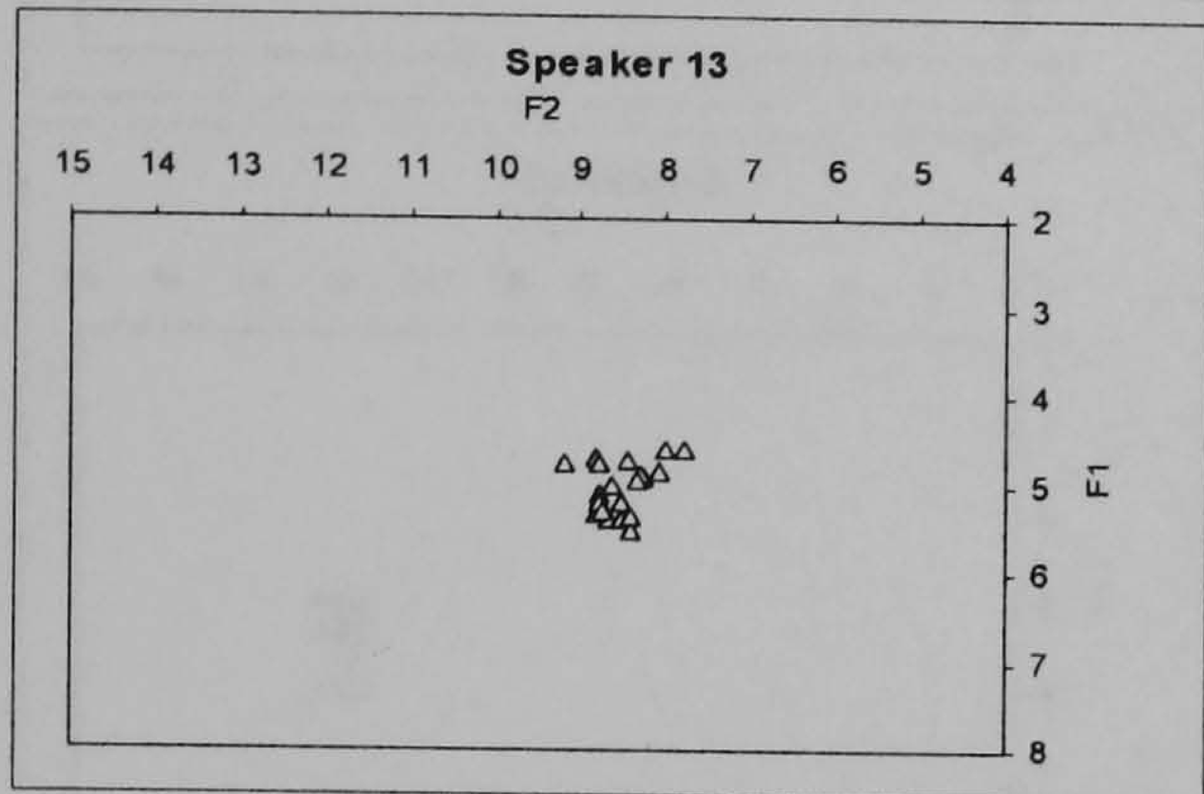

Speaker 11
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1
Speaker 12
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1
Speaker 13
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1
Speaker 14
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1
Speaker 15
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1
Speaker 16
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1
Speaker 17
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1
Speaker 18
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1
Speaker 19
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1
Speaker 20
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

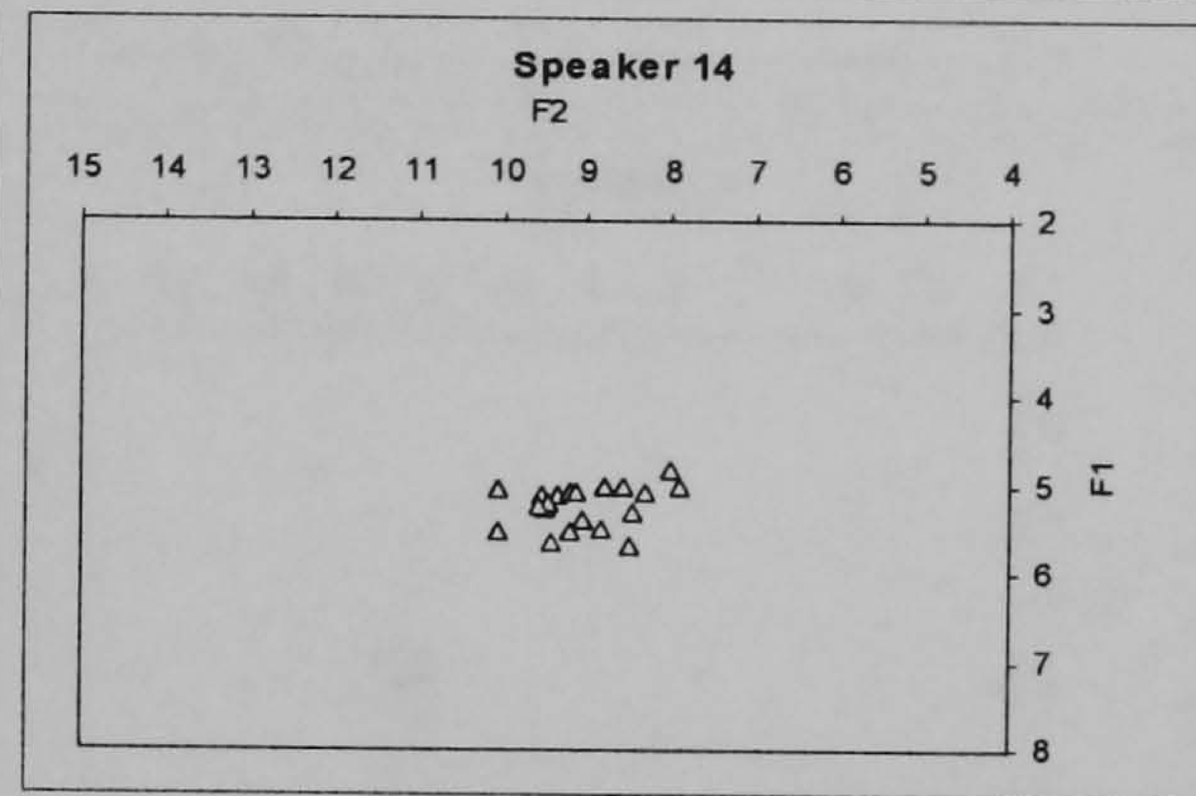
Production of /o:/ by speakers individually (Bark)

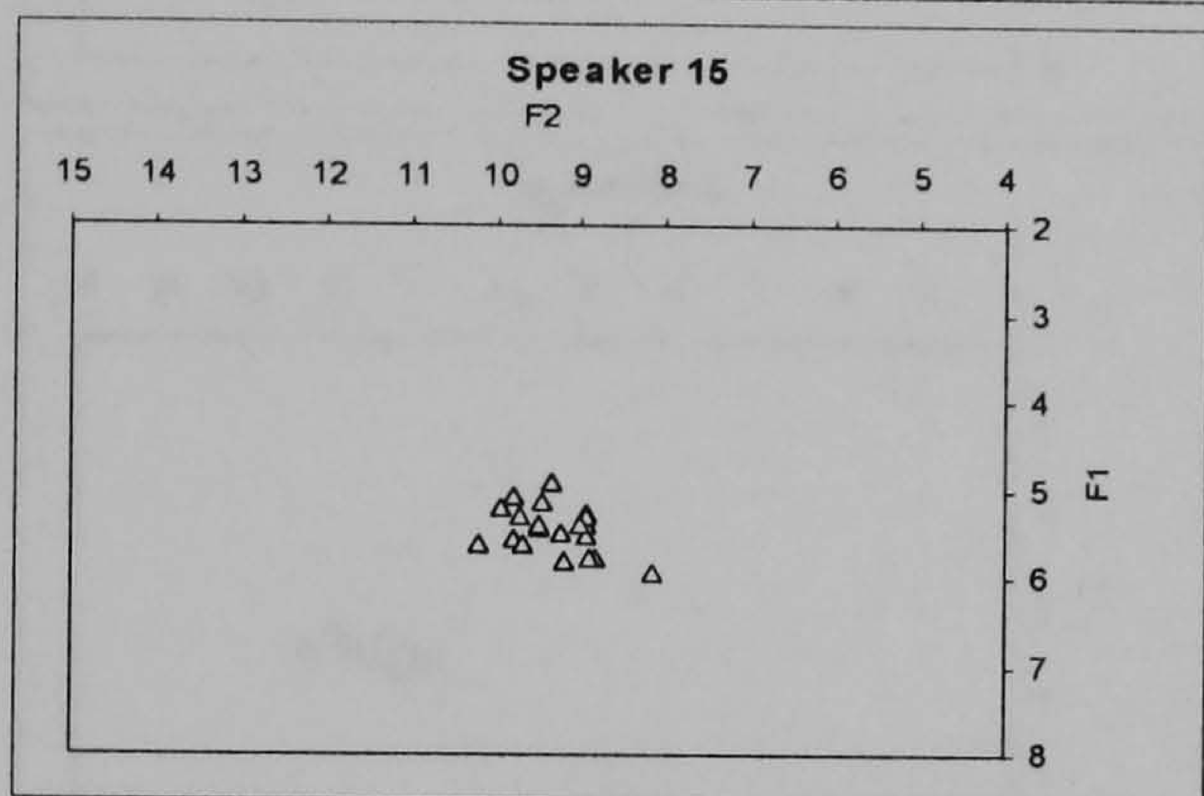


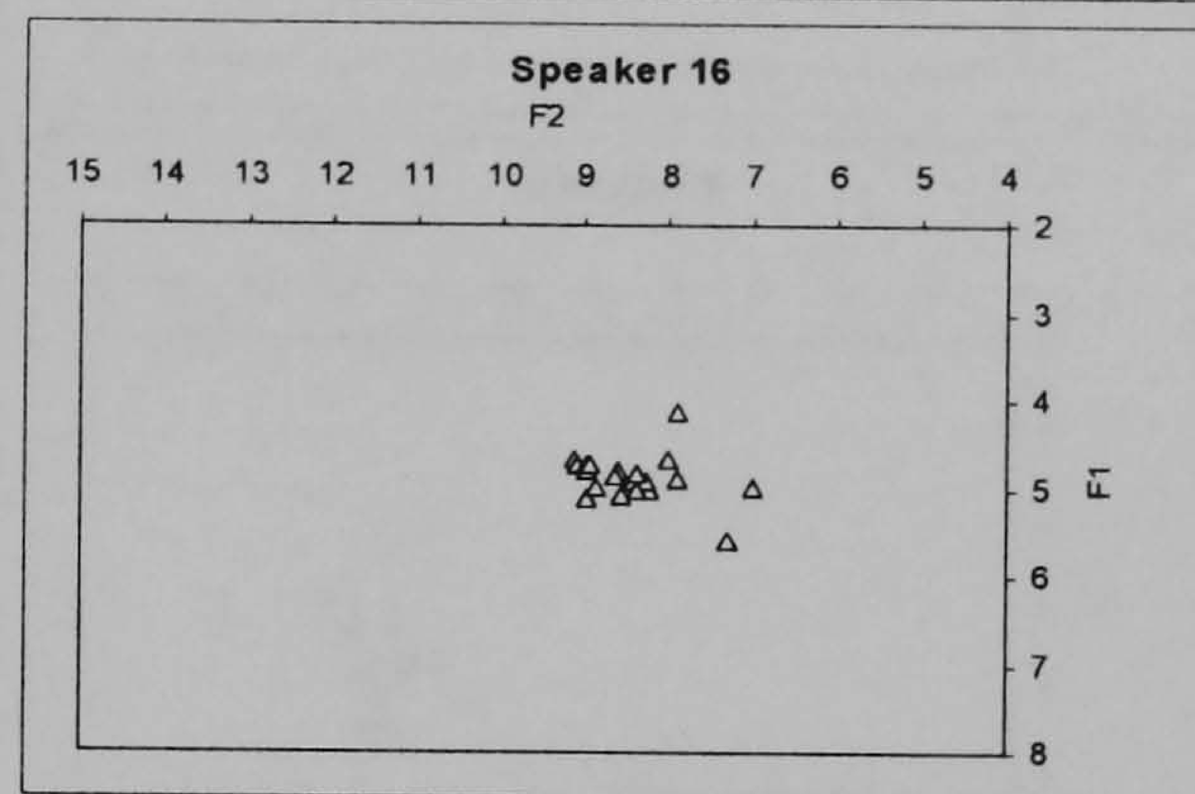

Speaker 11
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

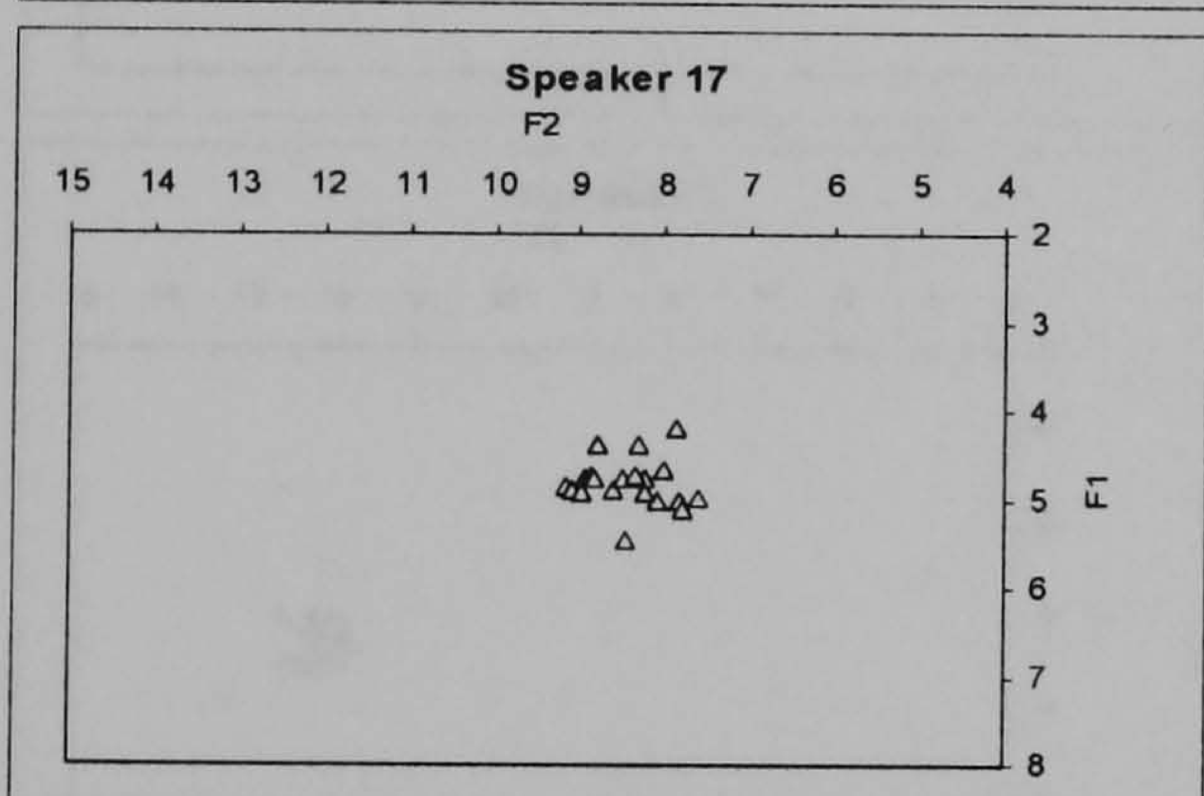

Speaker 12
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

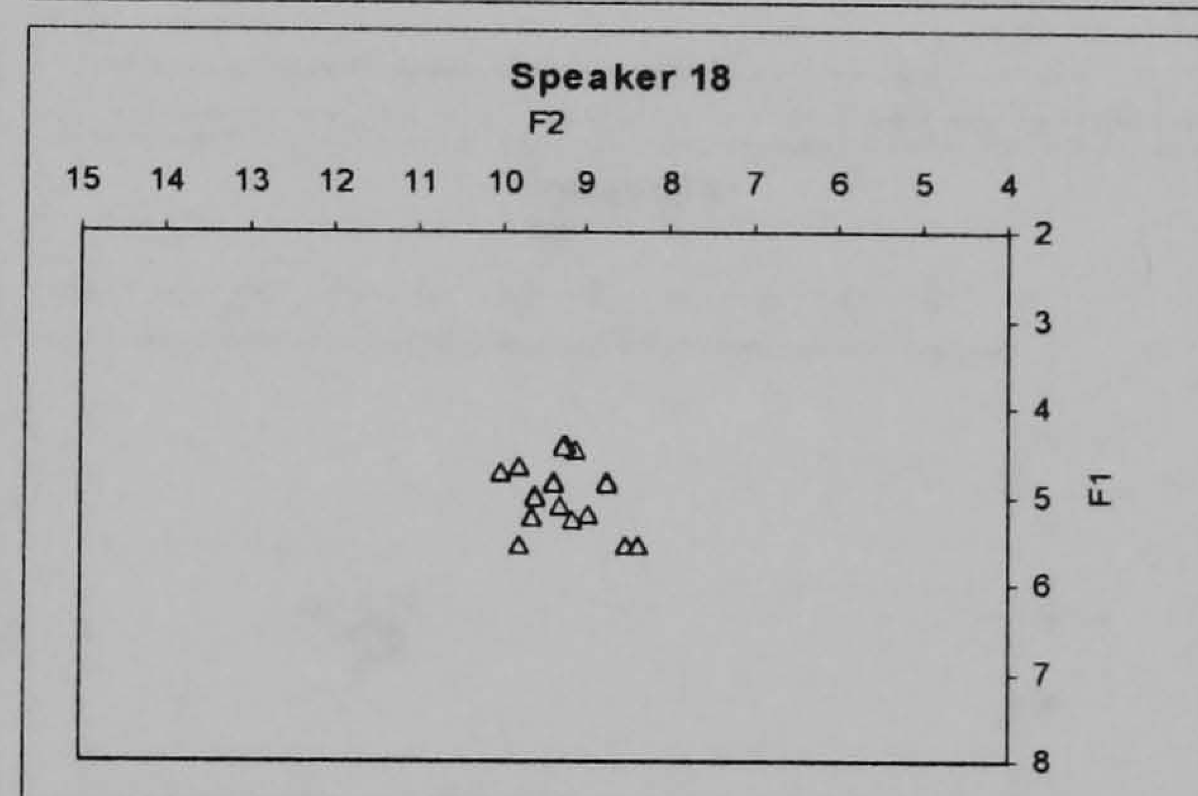

Speaker 13
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

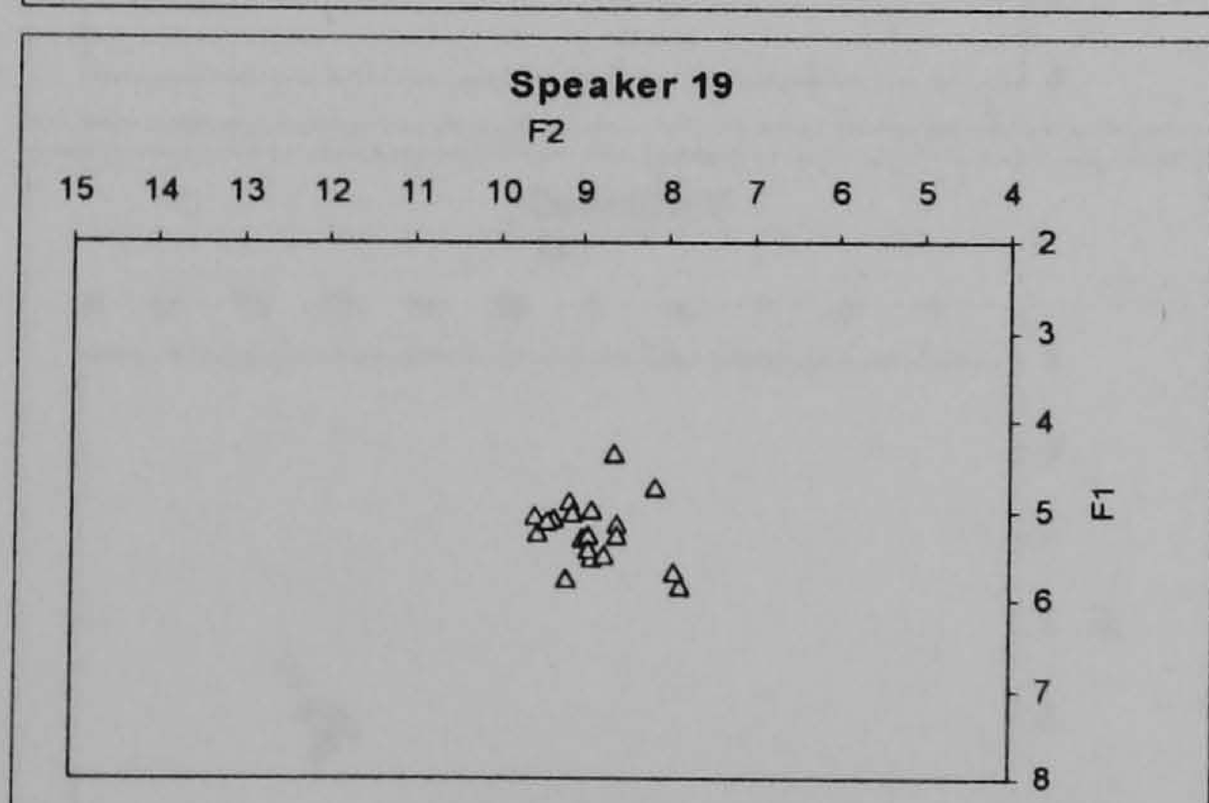

Speaker 14
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

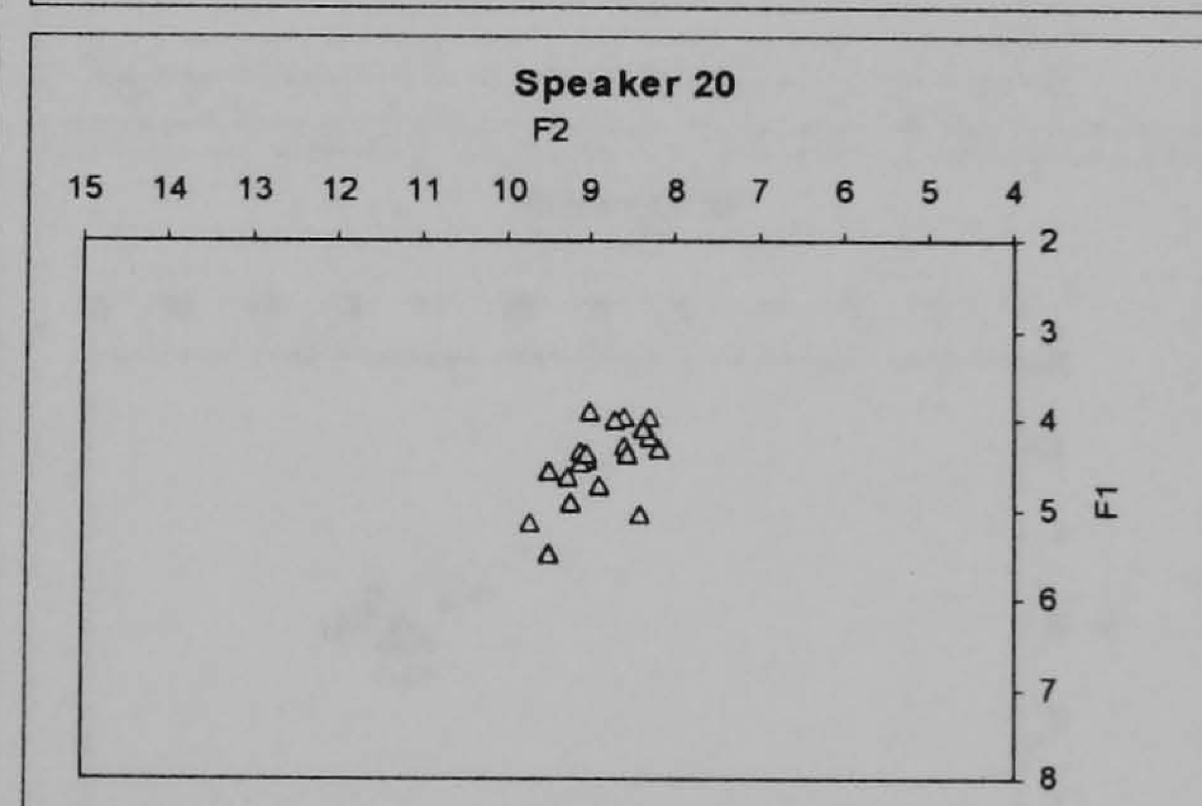

Speaker 15
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1


Speaker 16
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

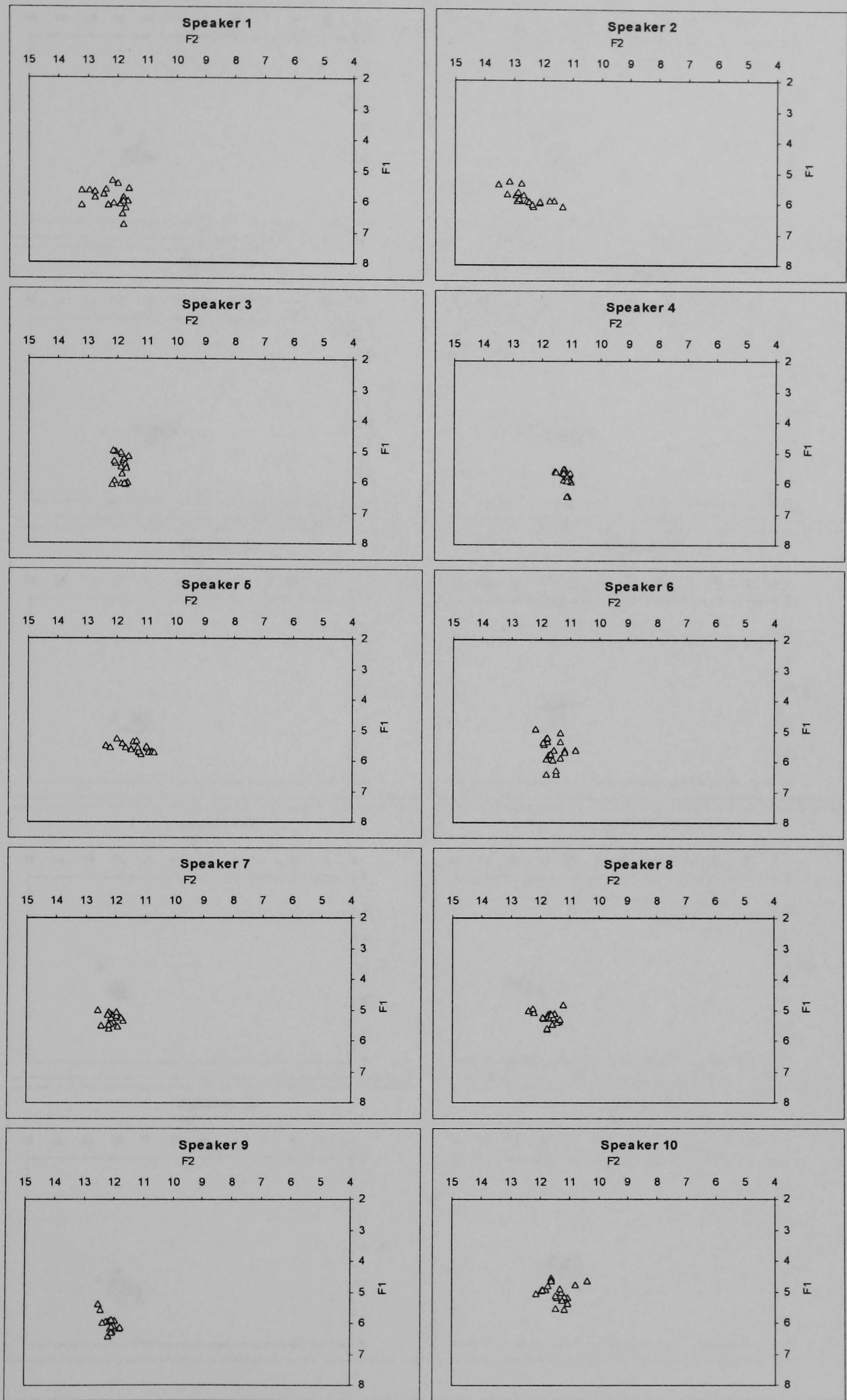

Speaker 17
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

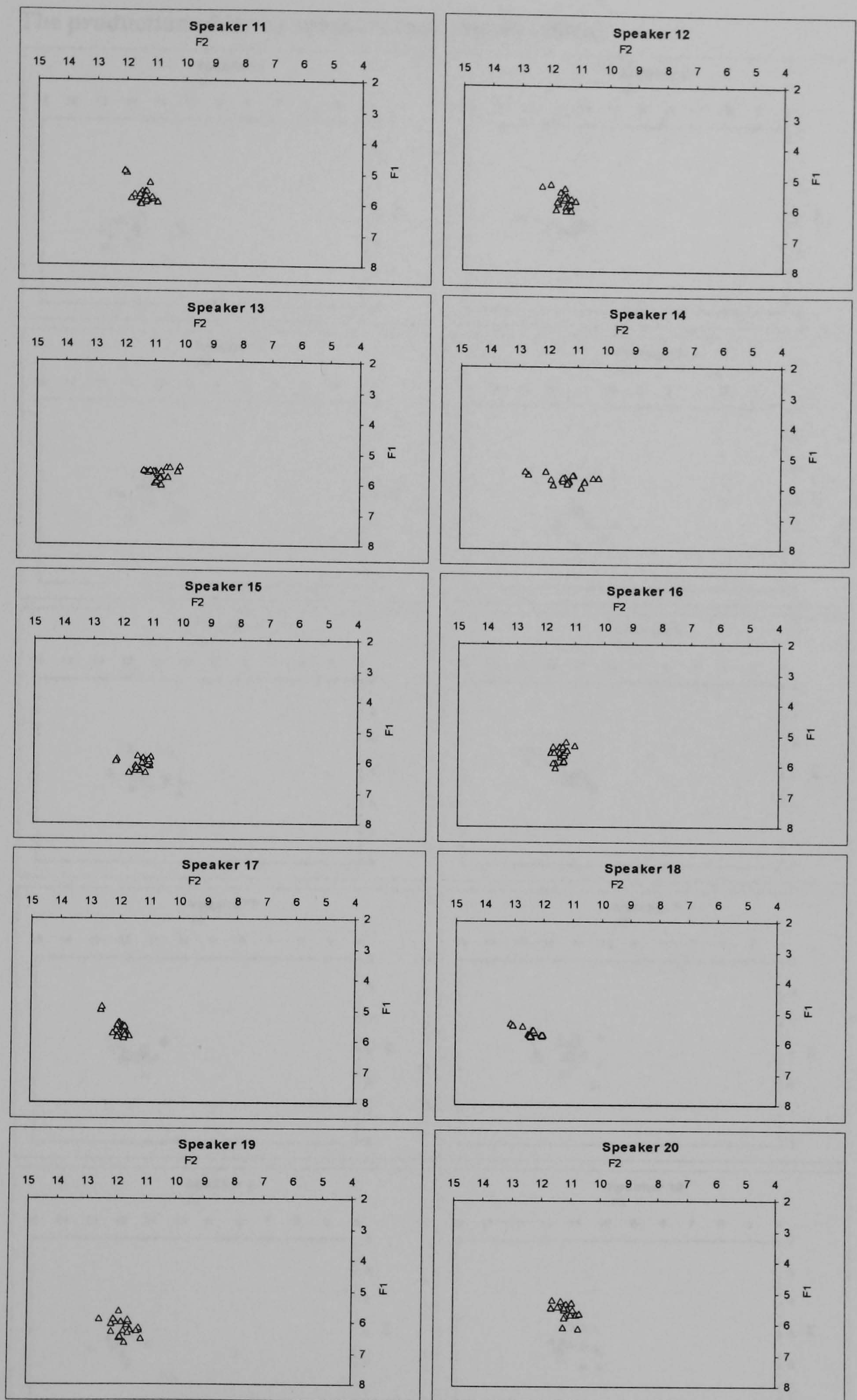

Speaker 18
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1


Speaker 19
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

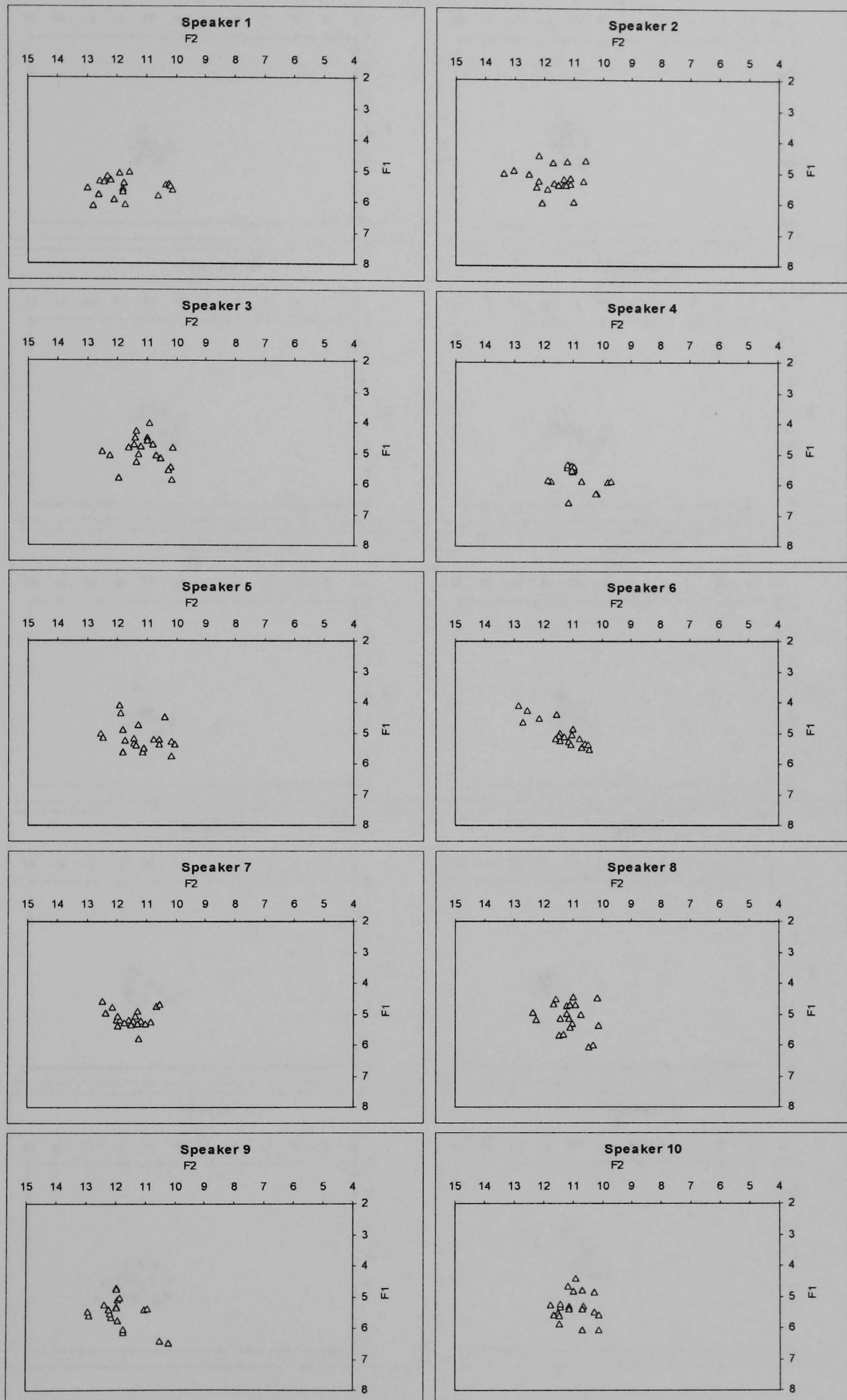

Speaker 20
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

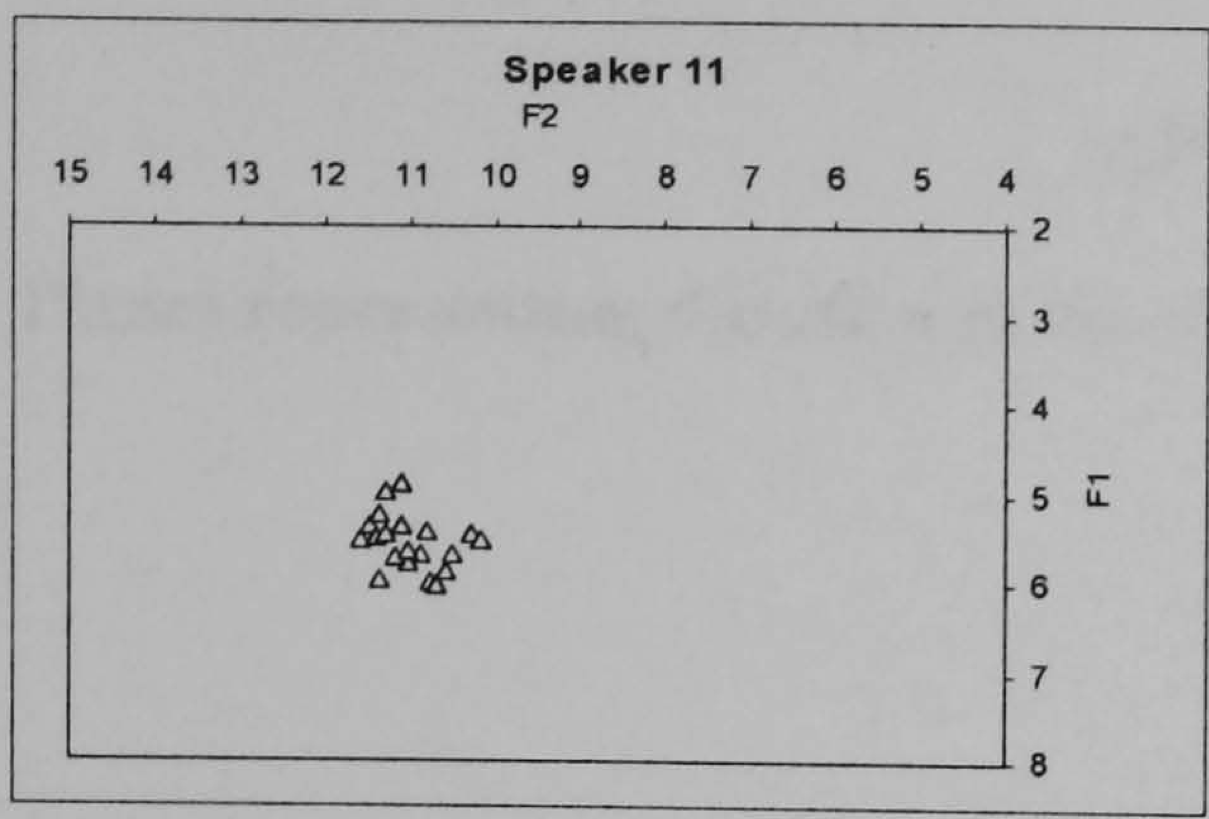
Production of /æ:/ by speakers individually (Bark)

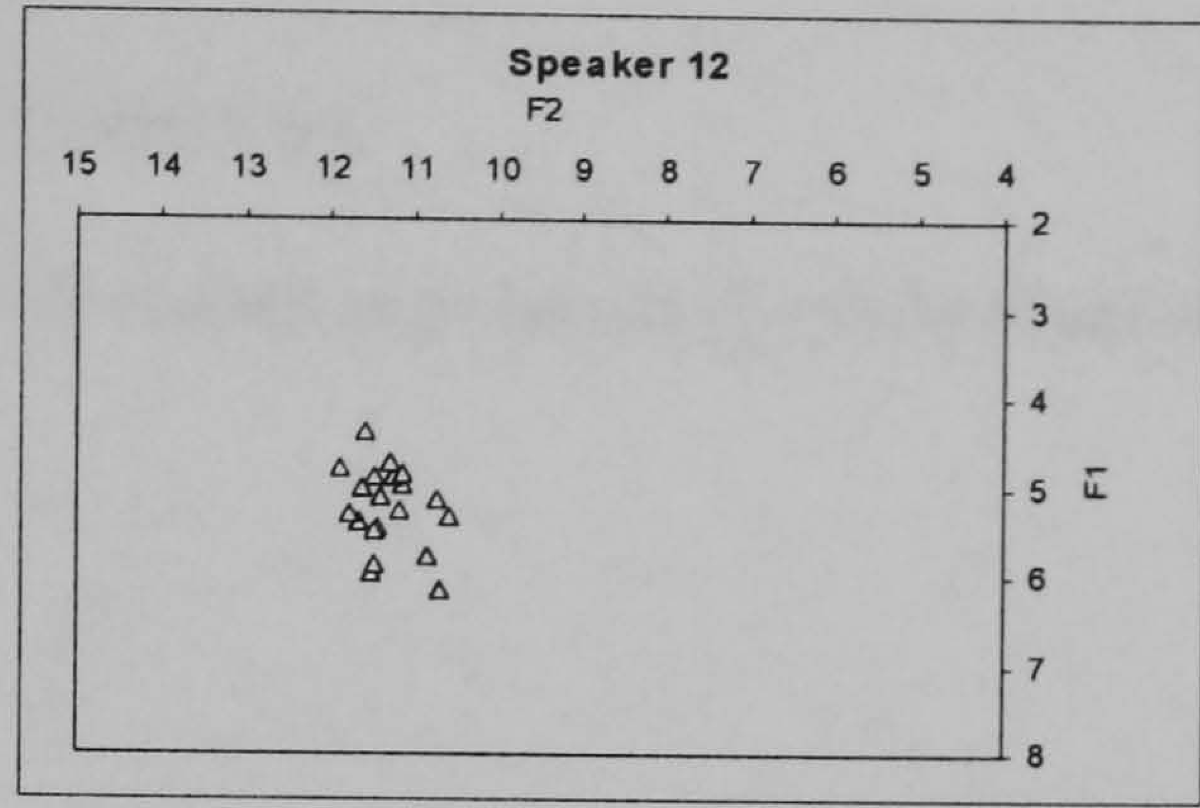


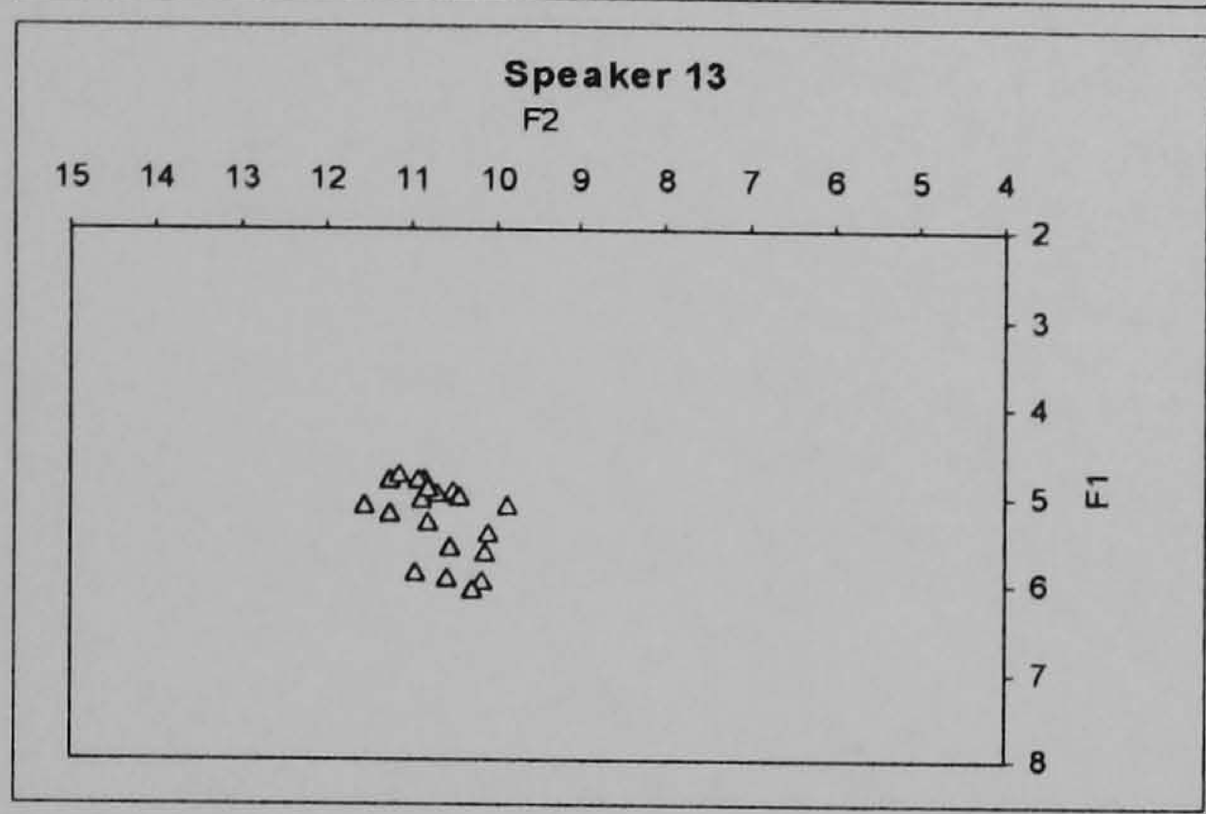

Speaker 11
F2
F1
Speaker 12
F2
F1
Speaker 13
F2
F1
Speaker 14
F2
F1
Speaker 15
F2
F1
Speaker 16
F2
F1
Speaker 17
F2
F1
Speaker 18
F2
F1
Speaker 19
F2
F1
Speaker 20
F2
F1
15 14 13 12 11 10 9 8 7 6 5 4
2
3
4
5
6
7
8

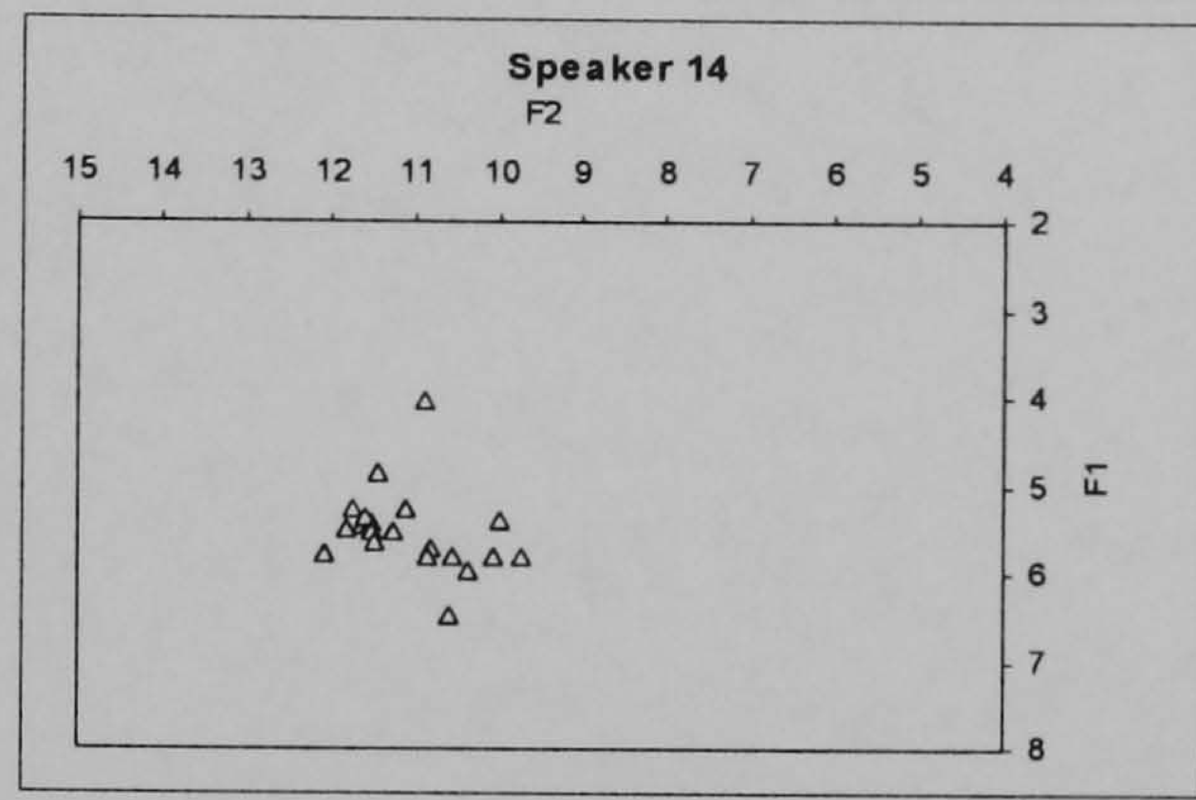
The production of /ə/ by speakers individually (Bark)

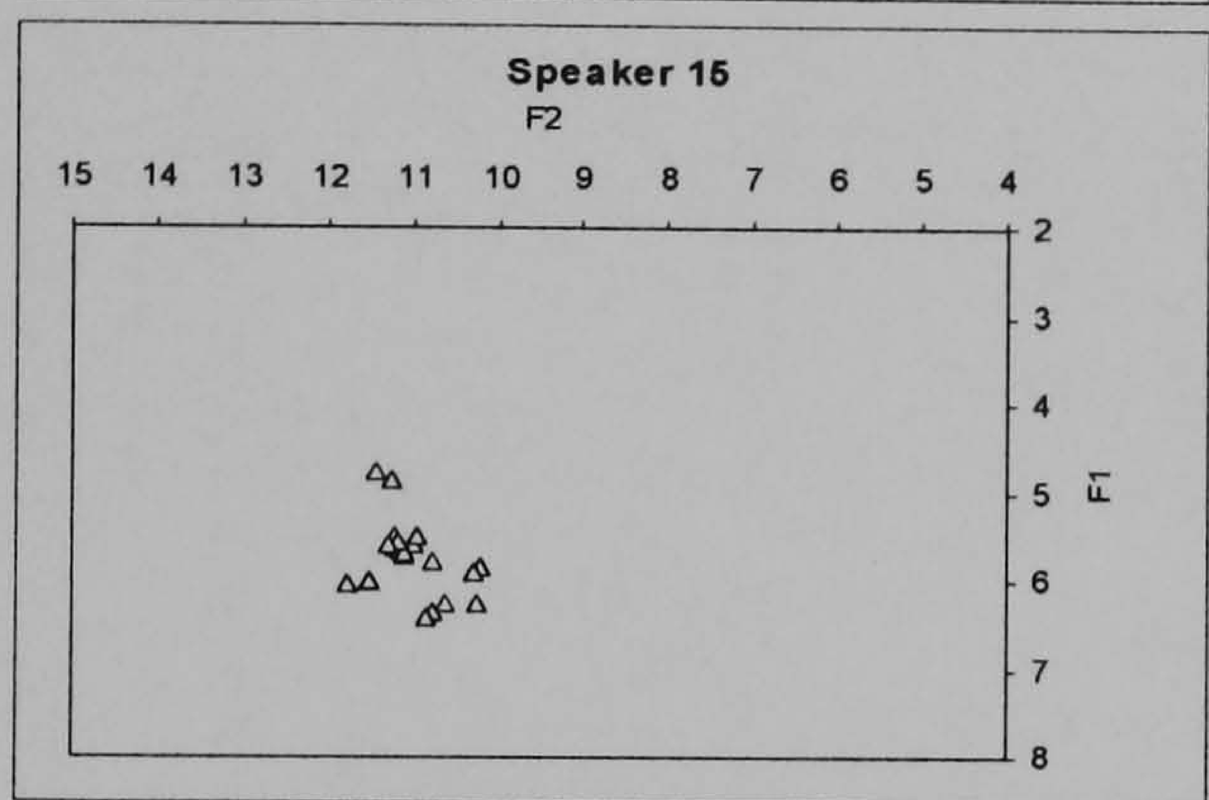


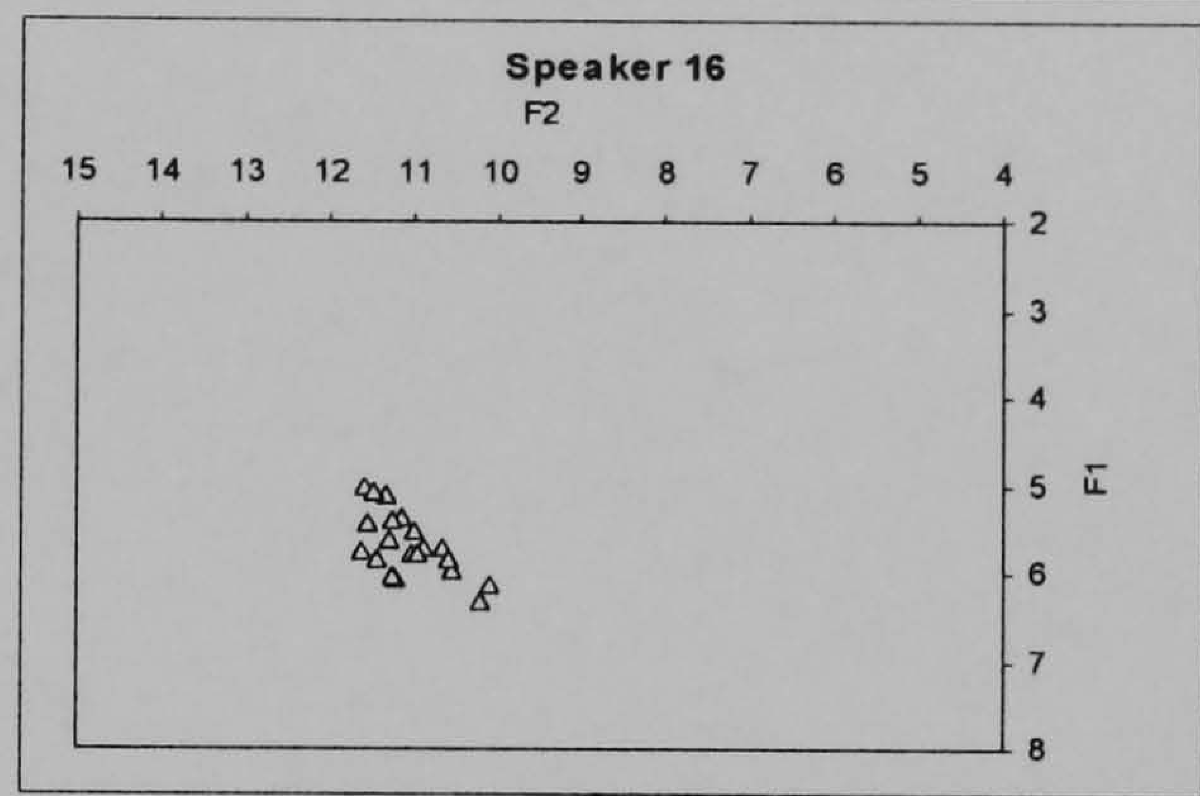

Speaker 11
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

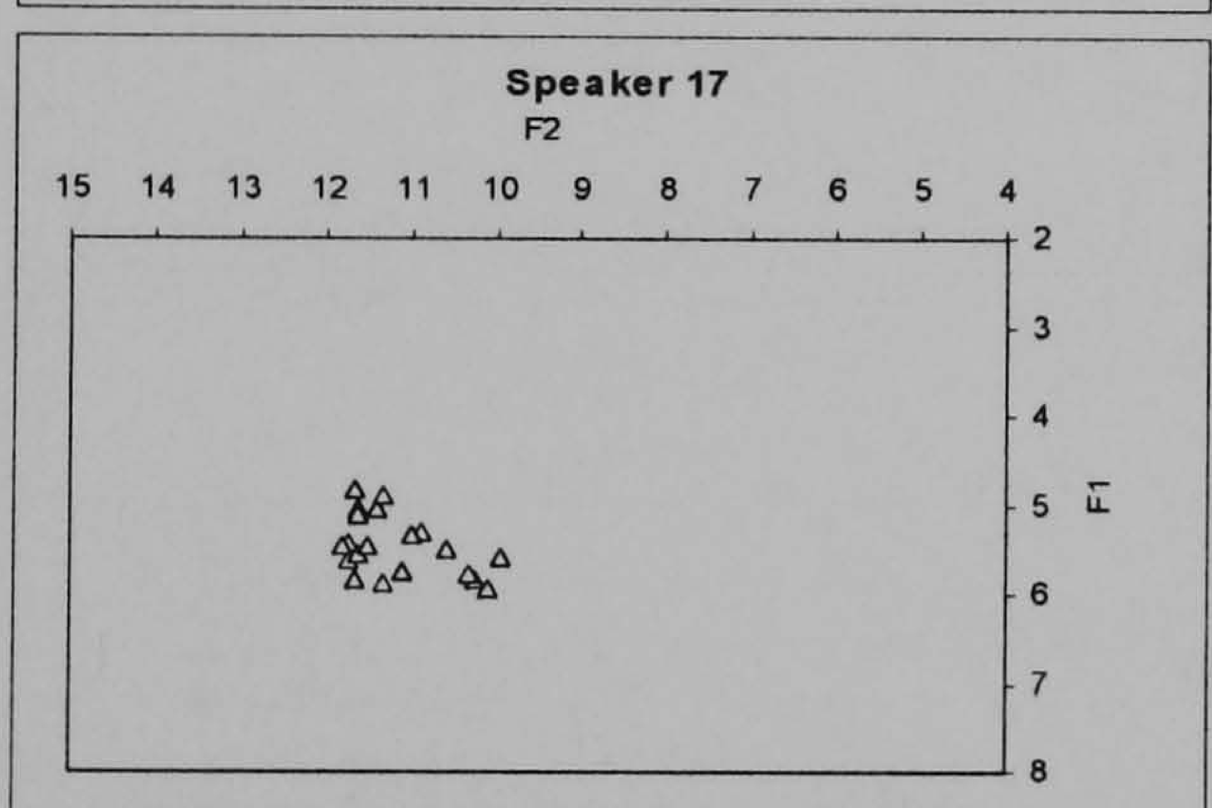

Speaker 12
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

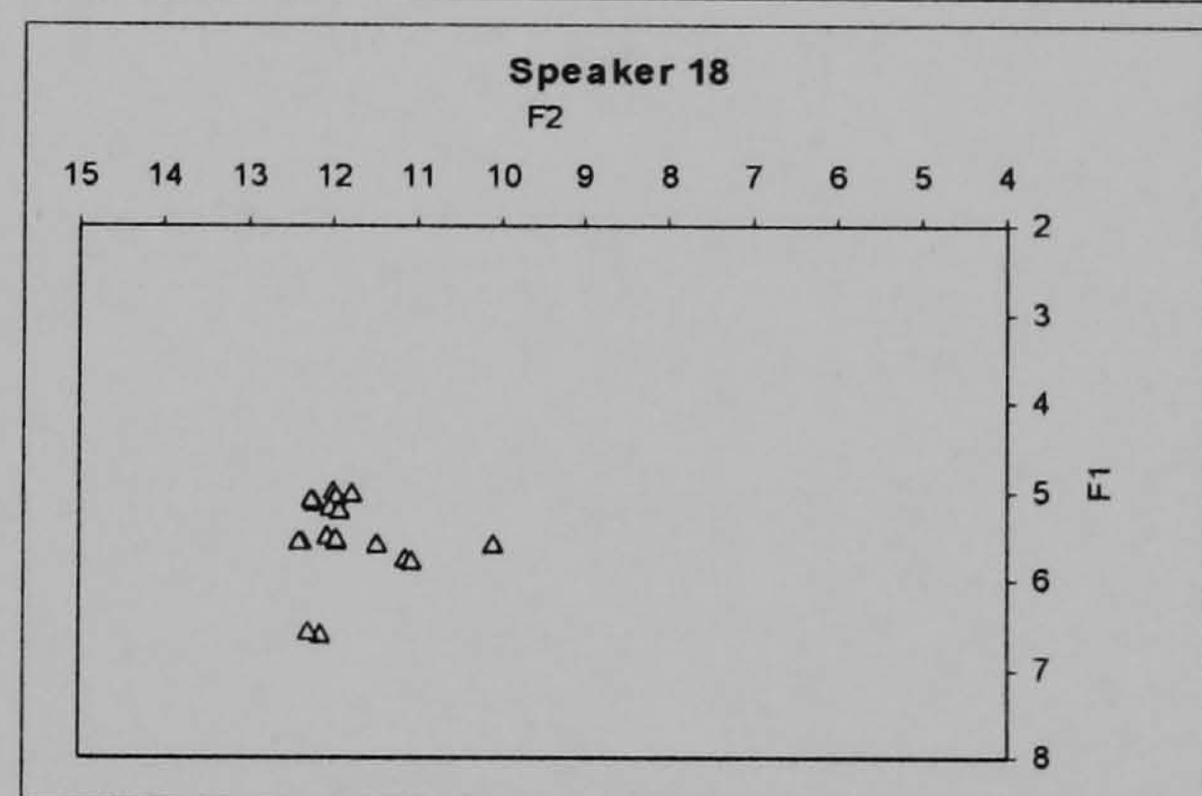

Speaker 13
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

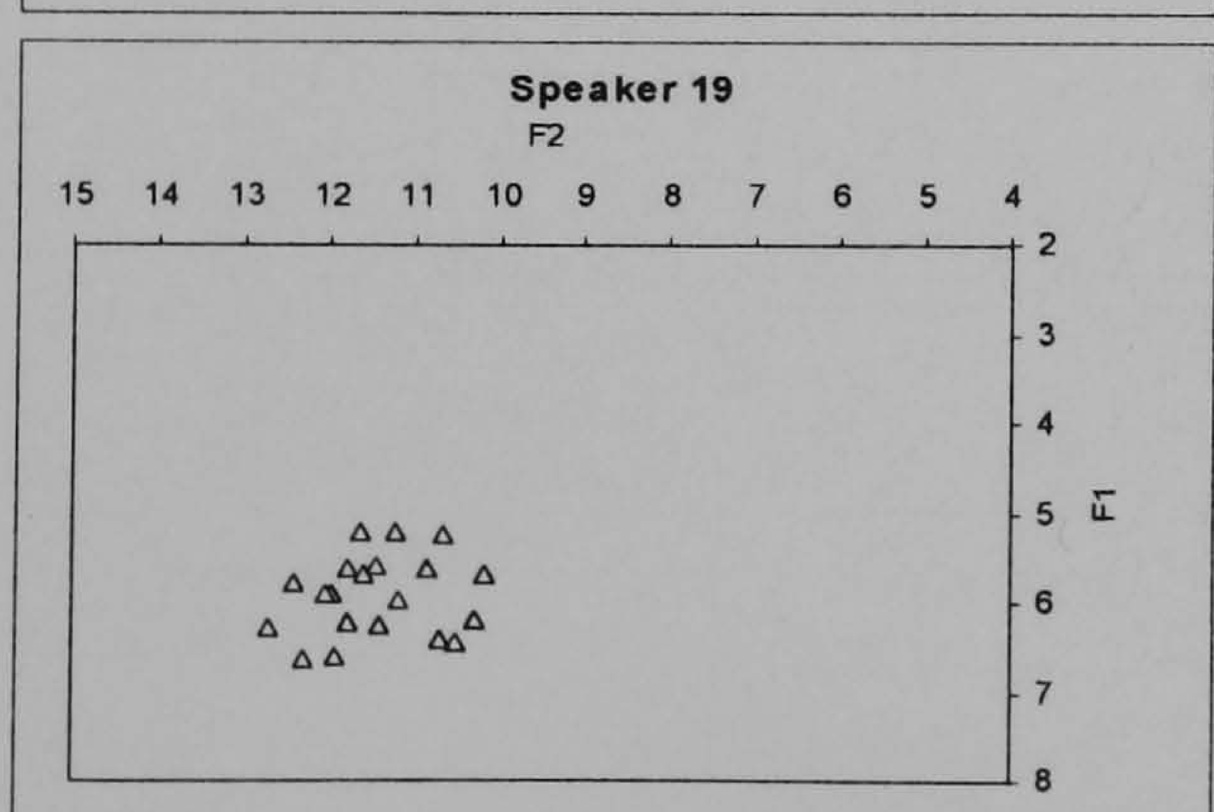

Speaker 14
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

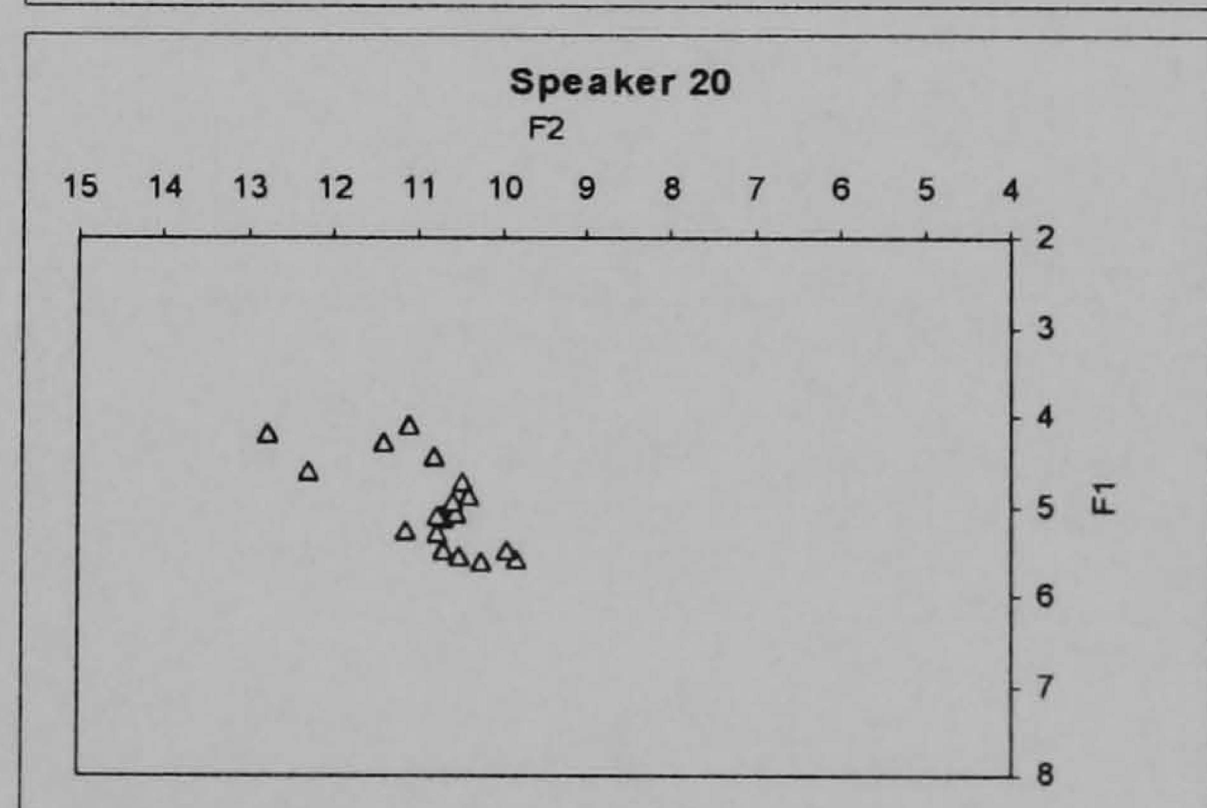

Speaker 15
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1


Speaker 16
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1


Speaker 17
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

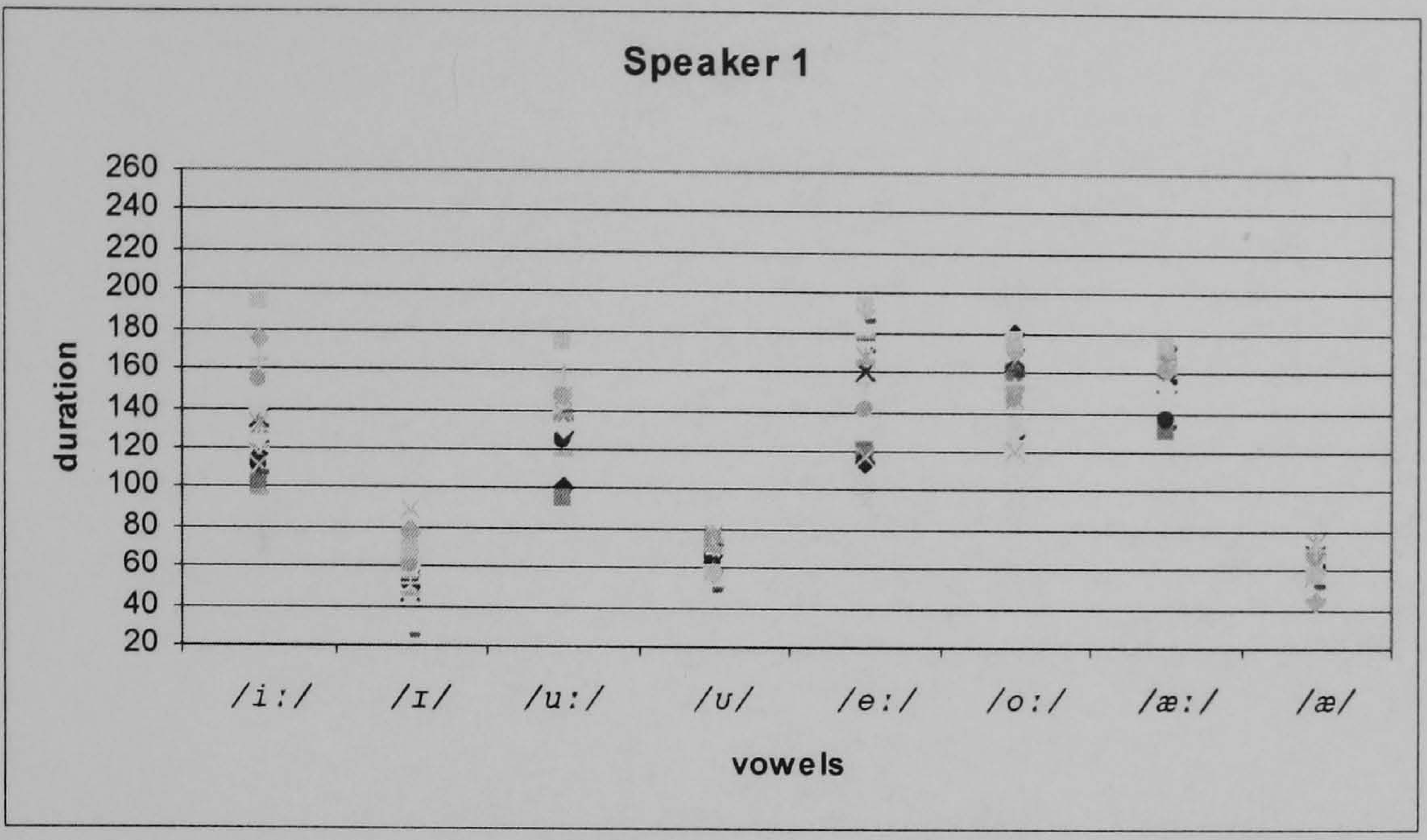

Speaker 18
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

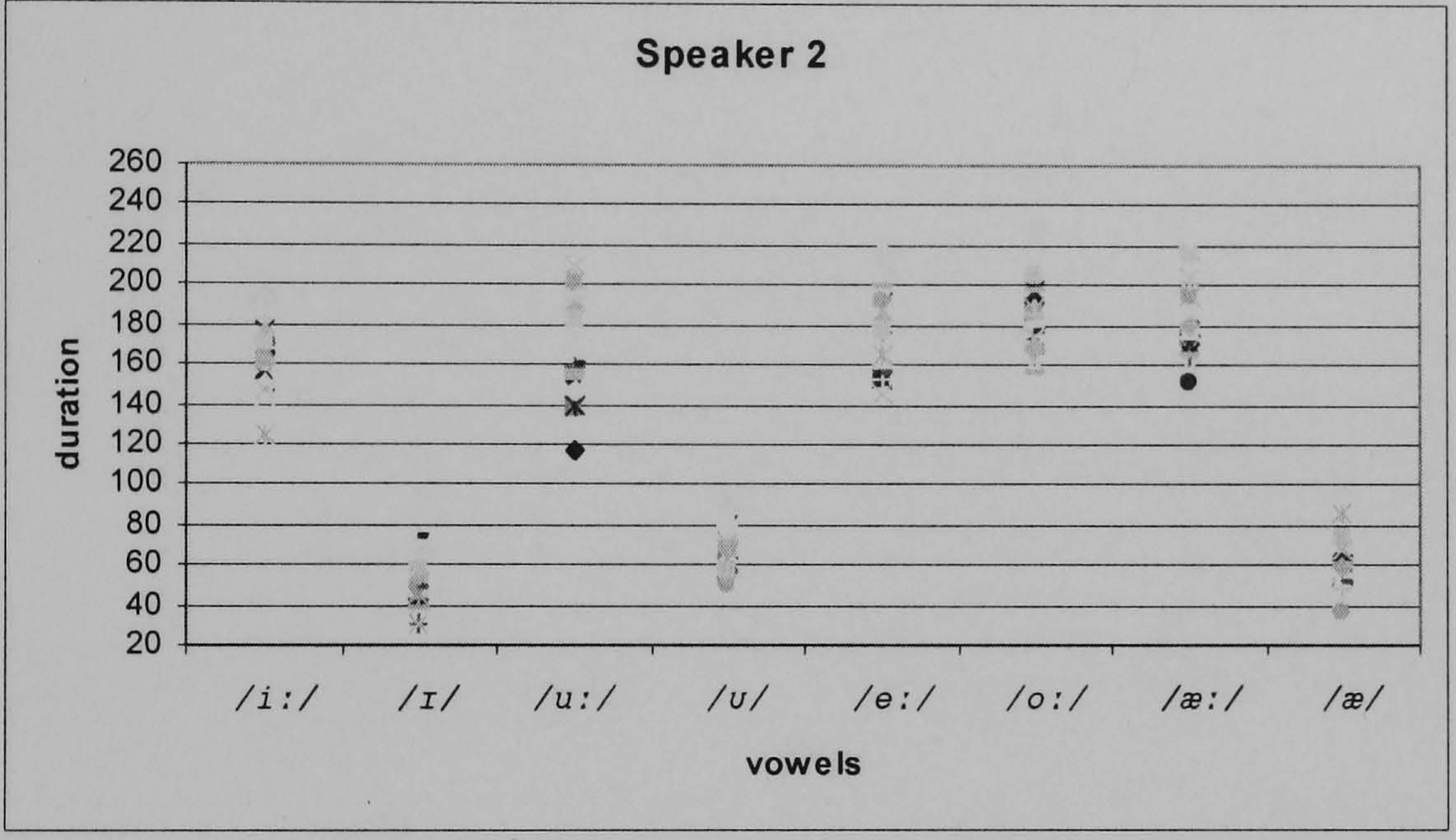

Speaker 19
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

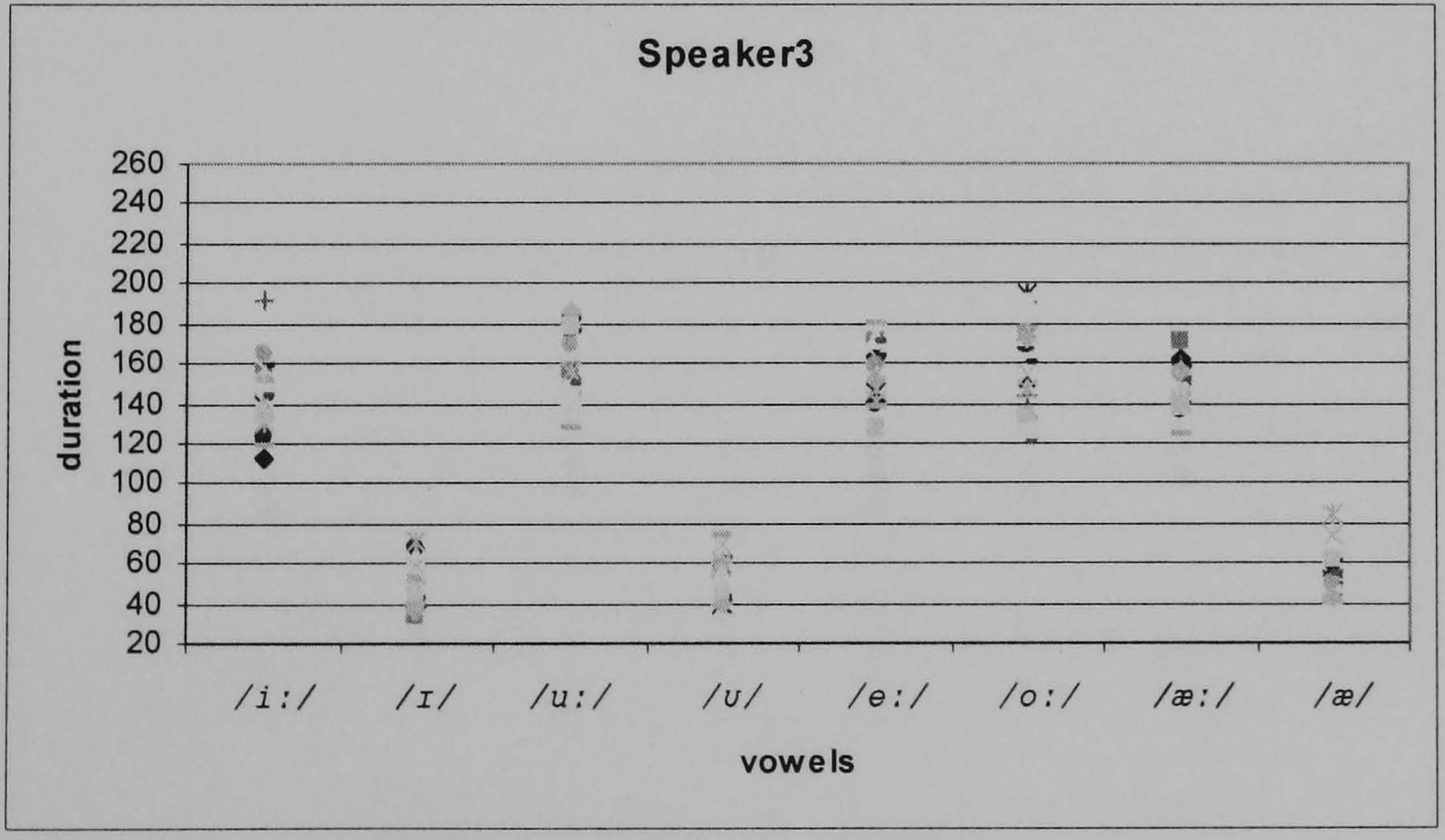

Speaker 20
F2
15 14 13 12 11 10 9 8 7 6 5 4
2 3 4 5 6 7 8
F1

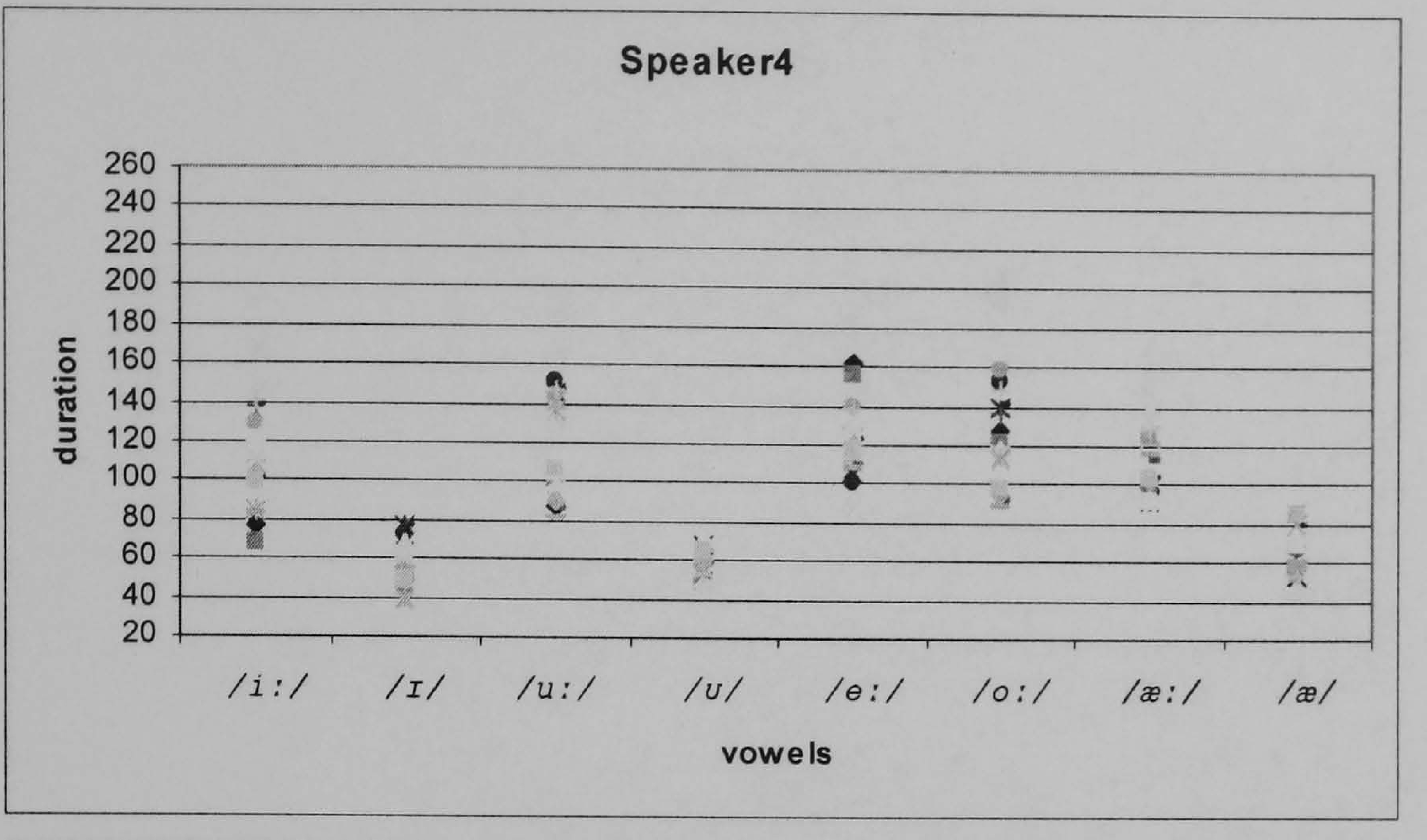
APPENDIX 3A

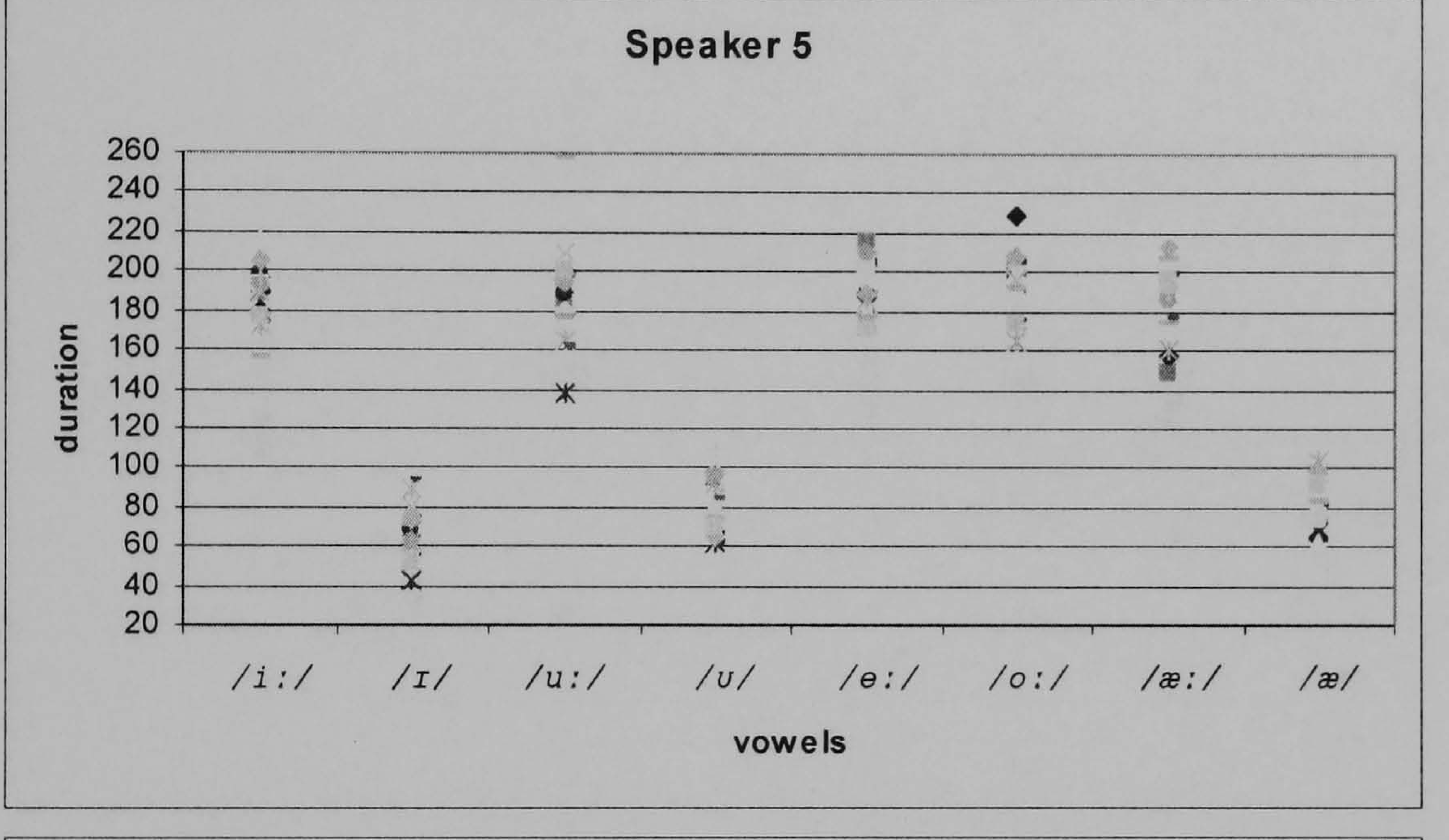
Planes representing duration in ms of all vowels as produced by individual speakers

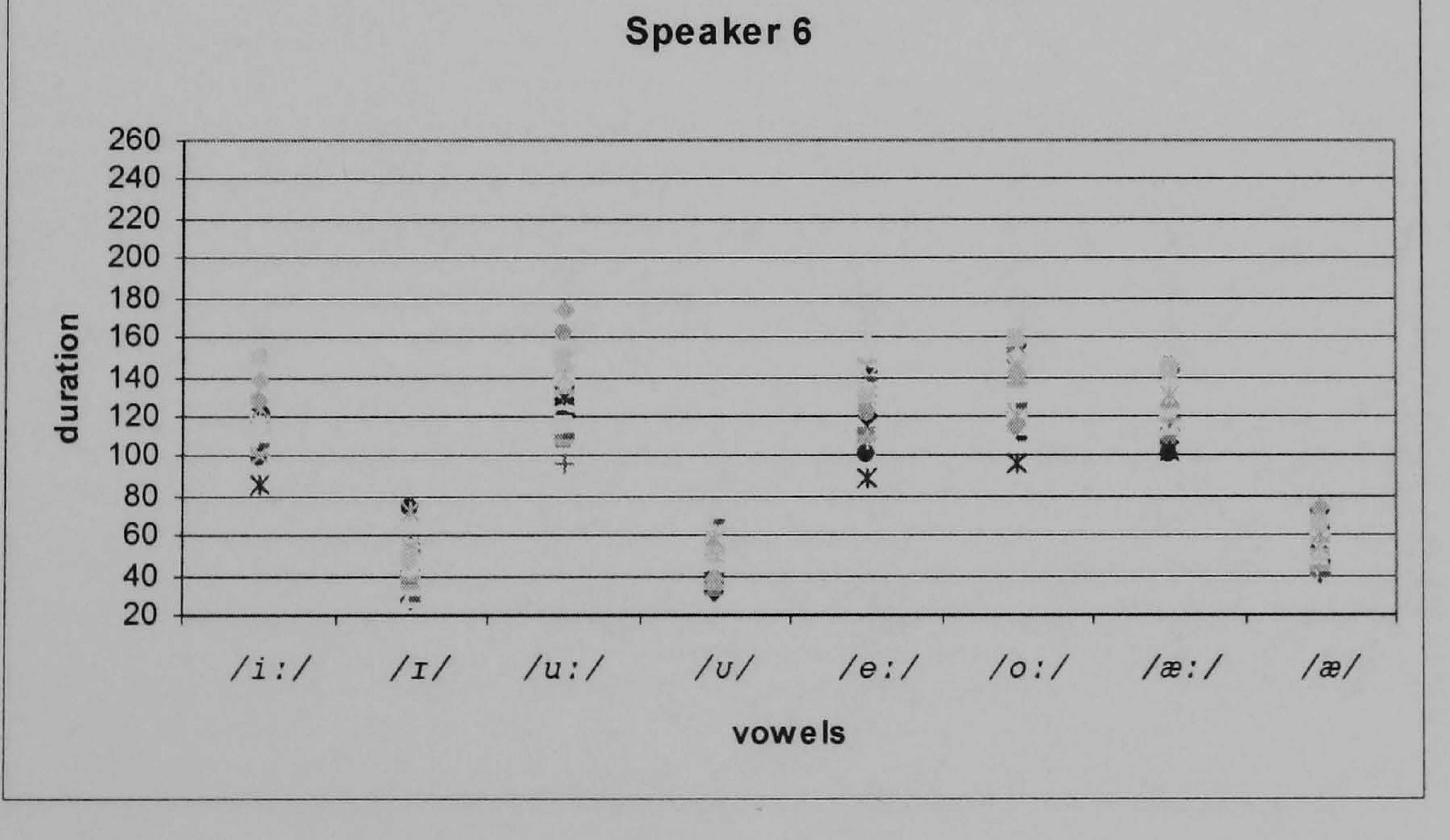

Speaker 1
260
240
220
200
180
160
140
120
100
80
60
40
20
duration
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels

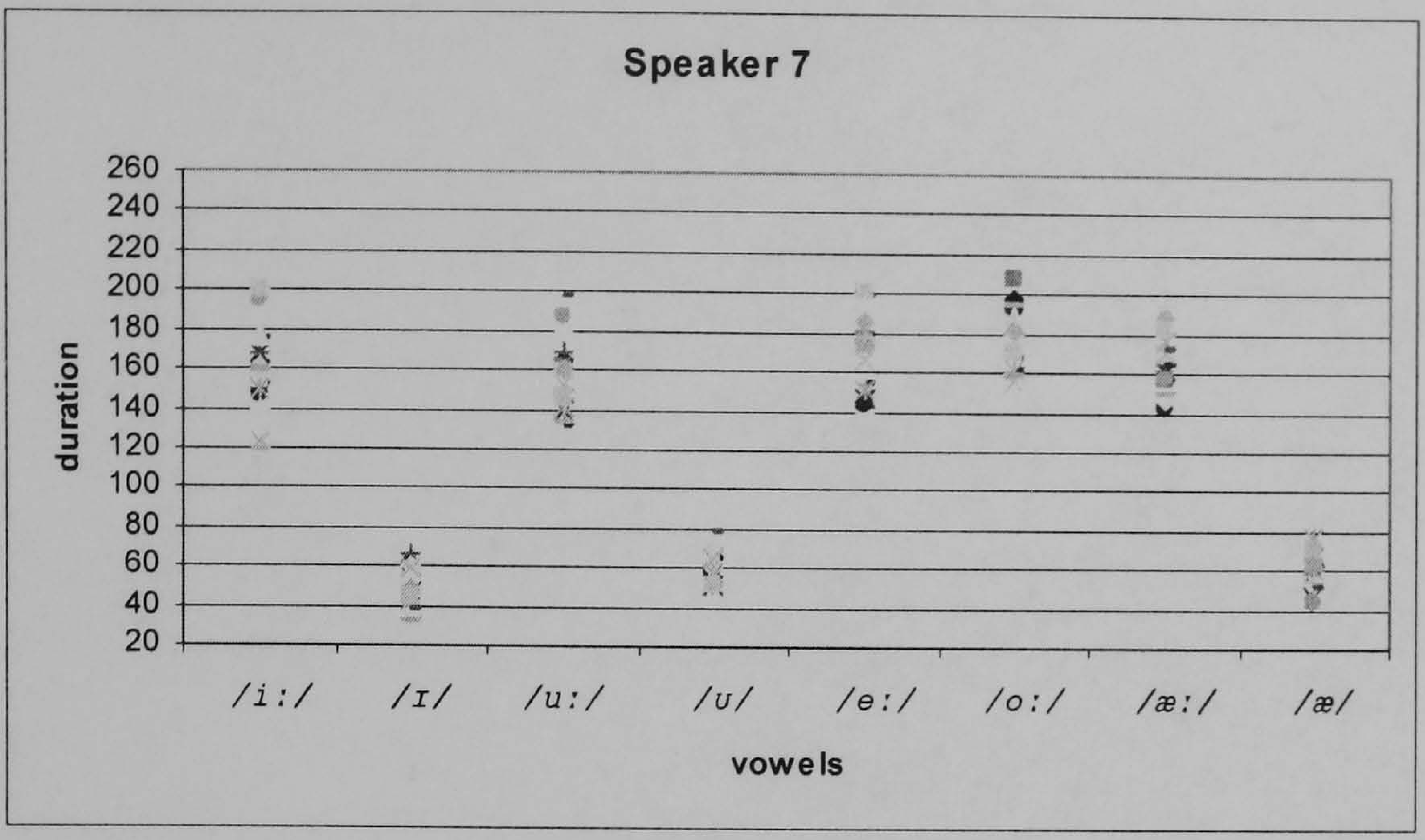

Speaker 2
260
240
220
200
180
160
140
120
100
80
60
40
20
duration
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels

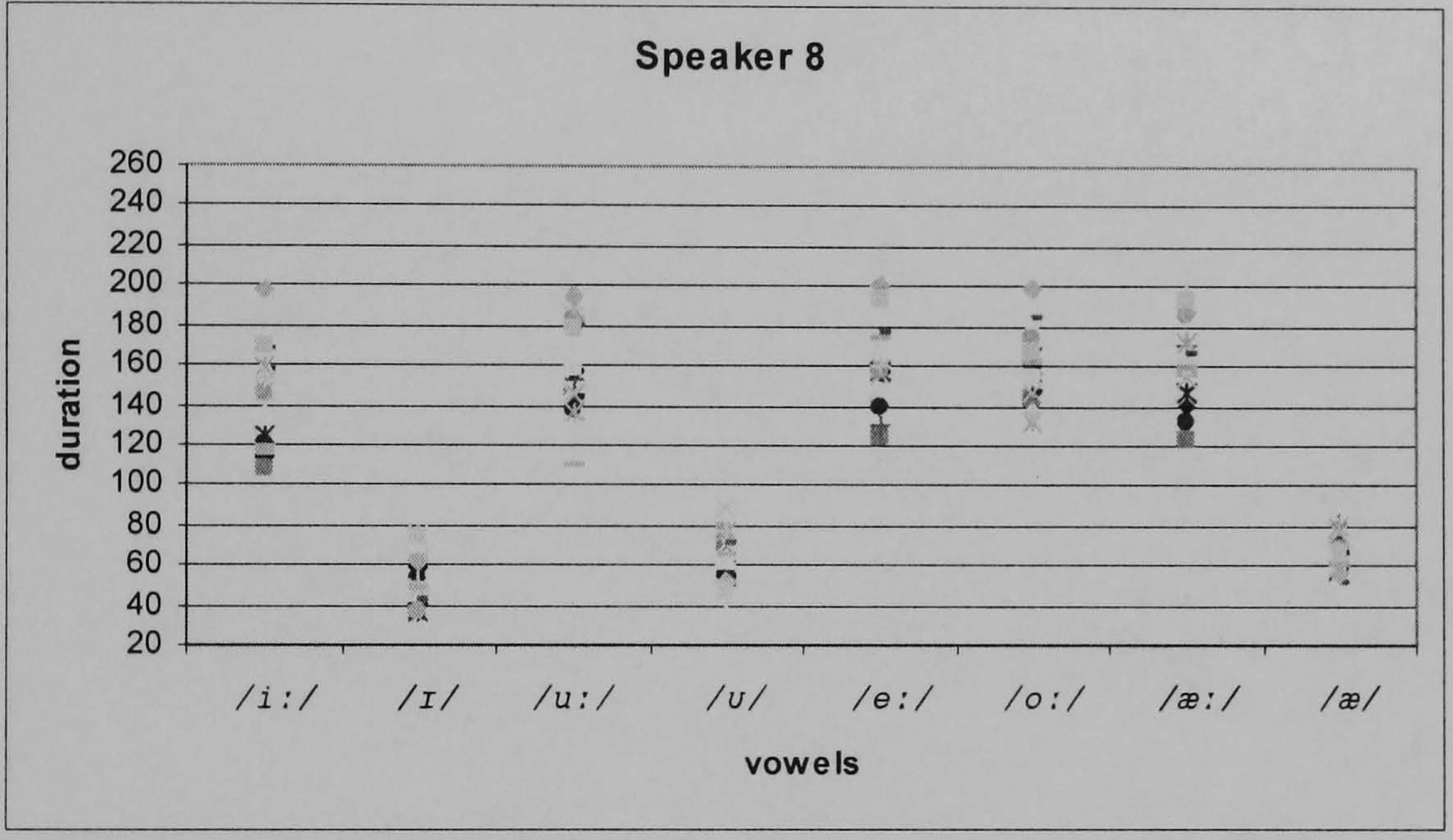

Speaker3
260
240
220
200
180
160
140
120
100
80
60
40
20
duration
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels

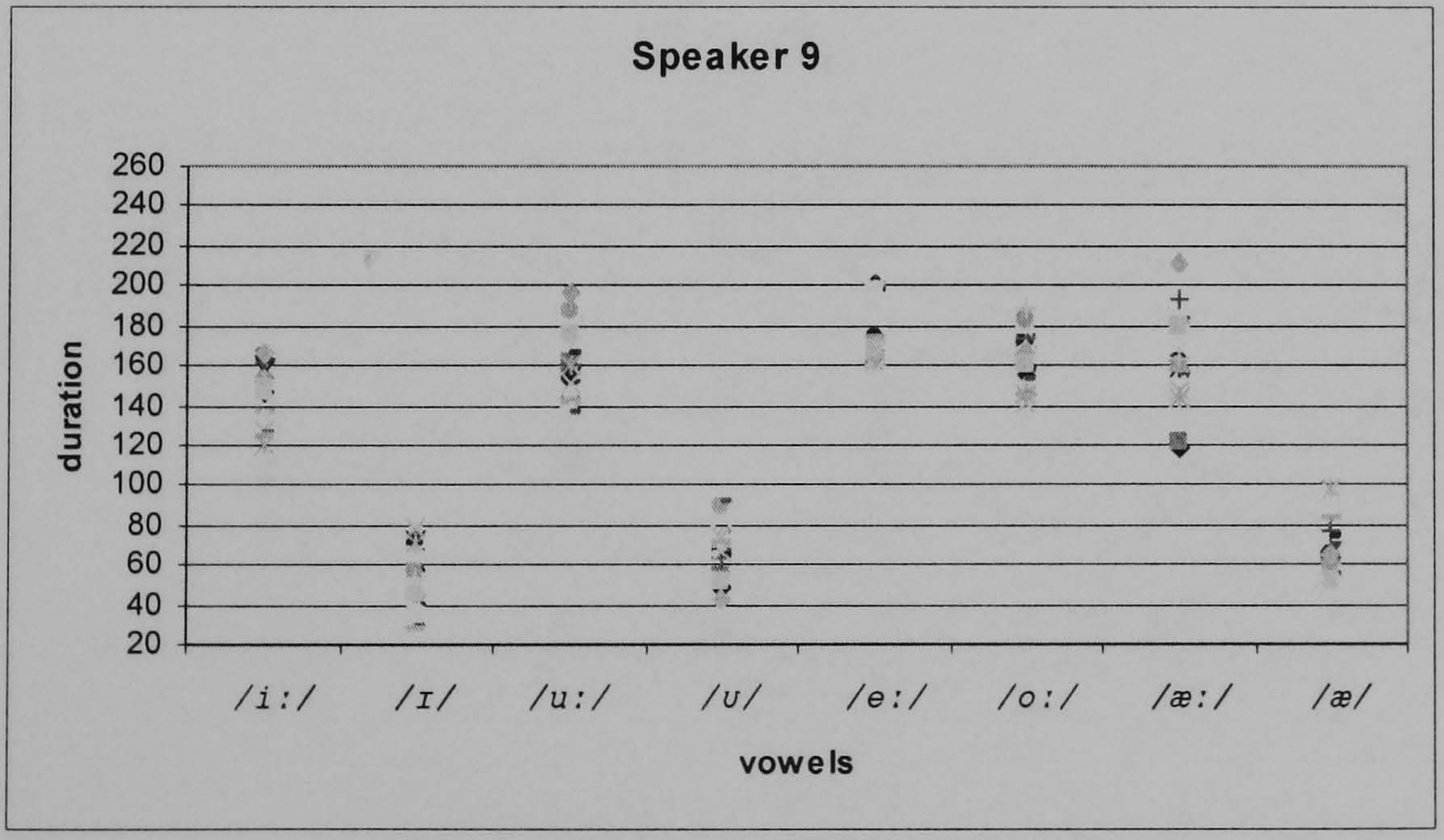

Speaker4
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels

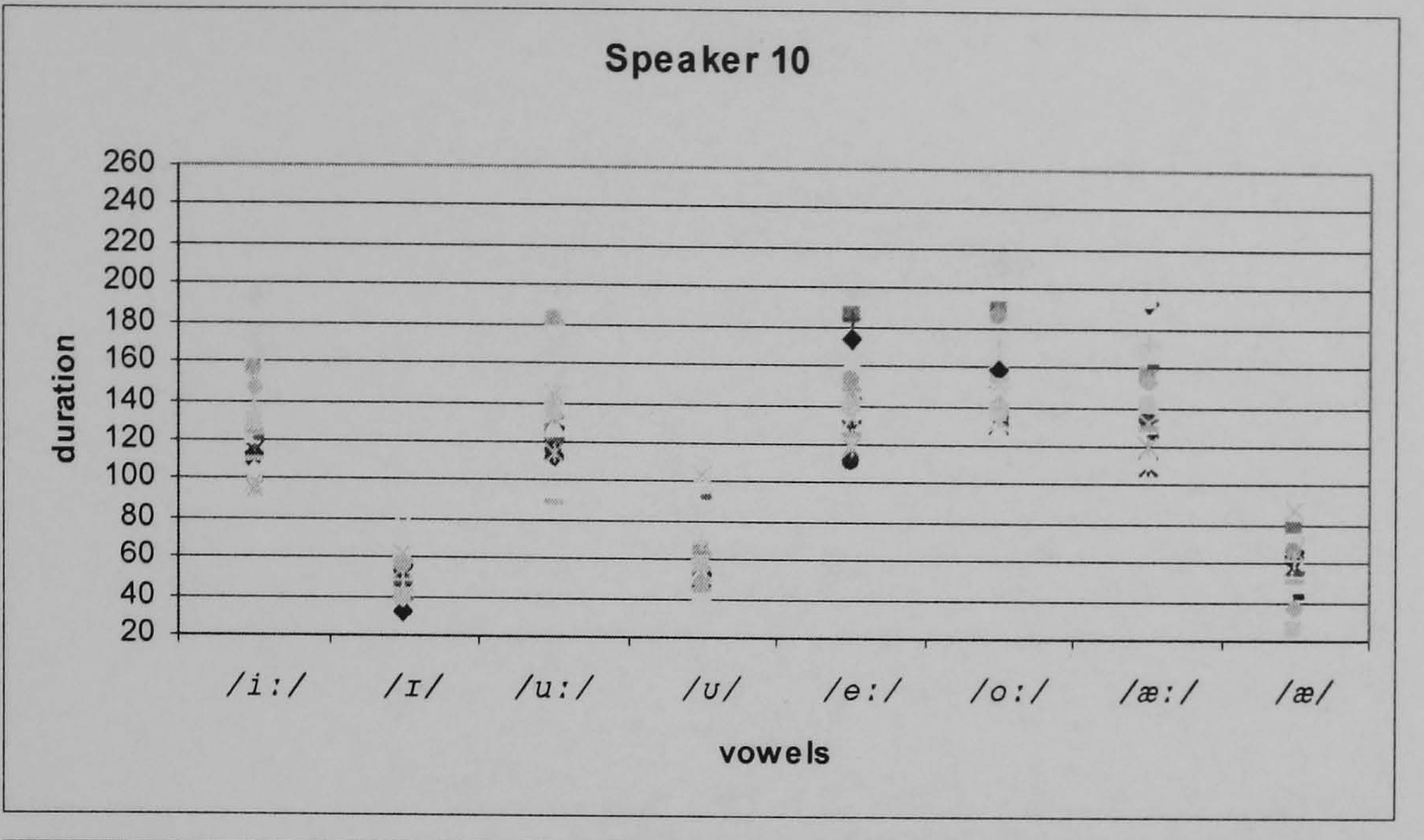

Speaker 5
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels

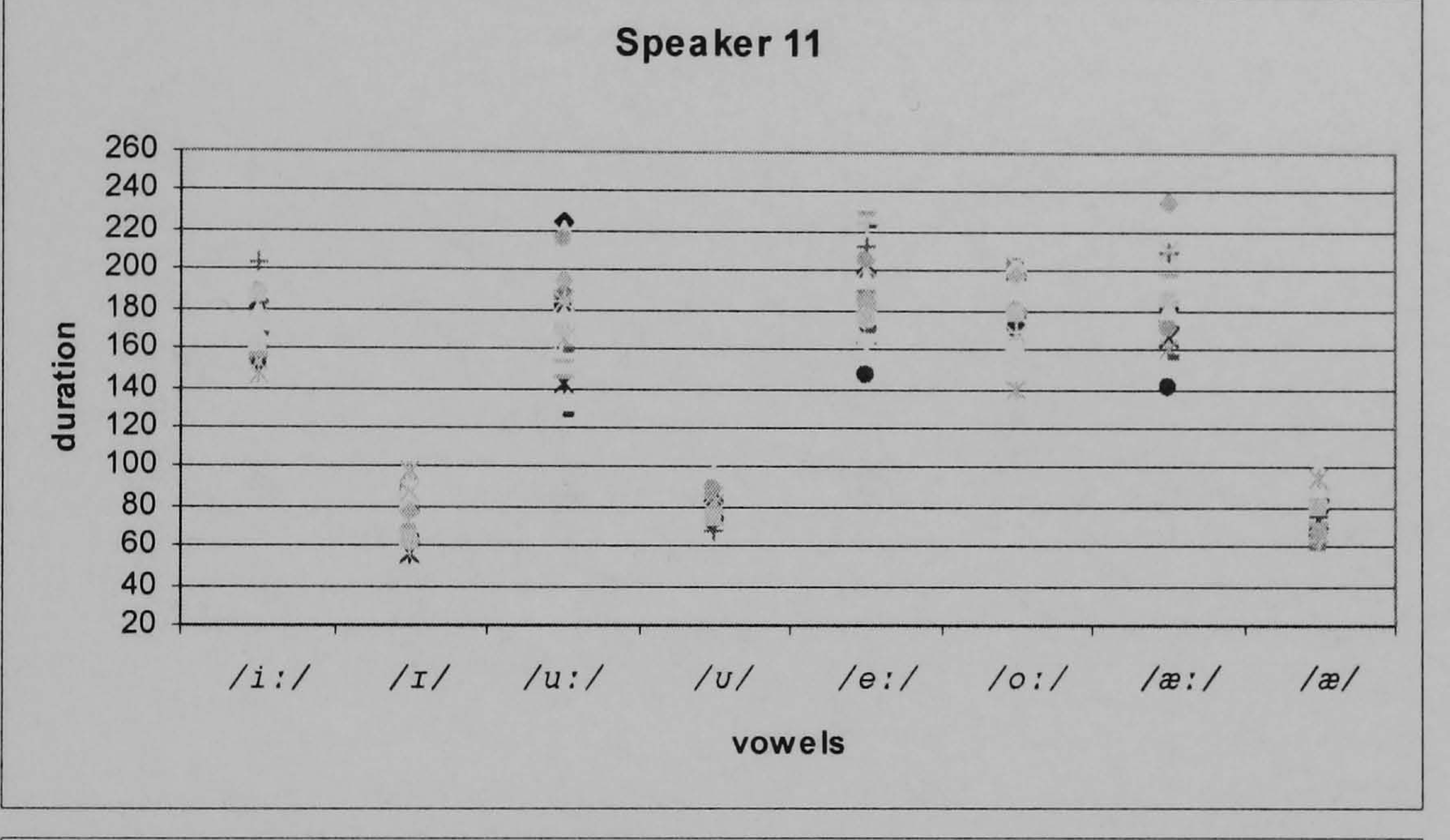

Speaker 6
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels

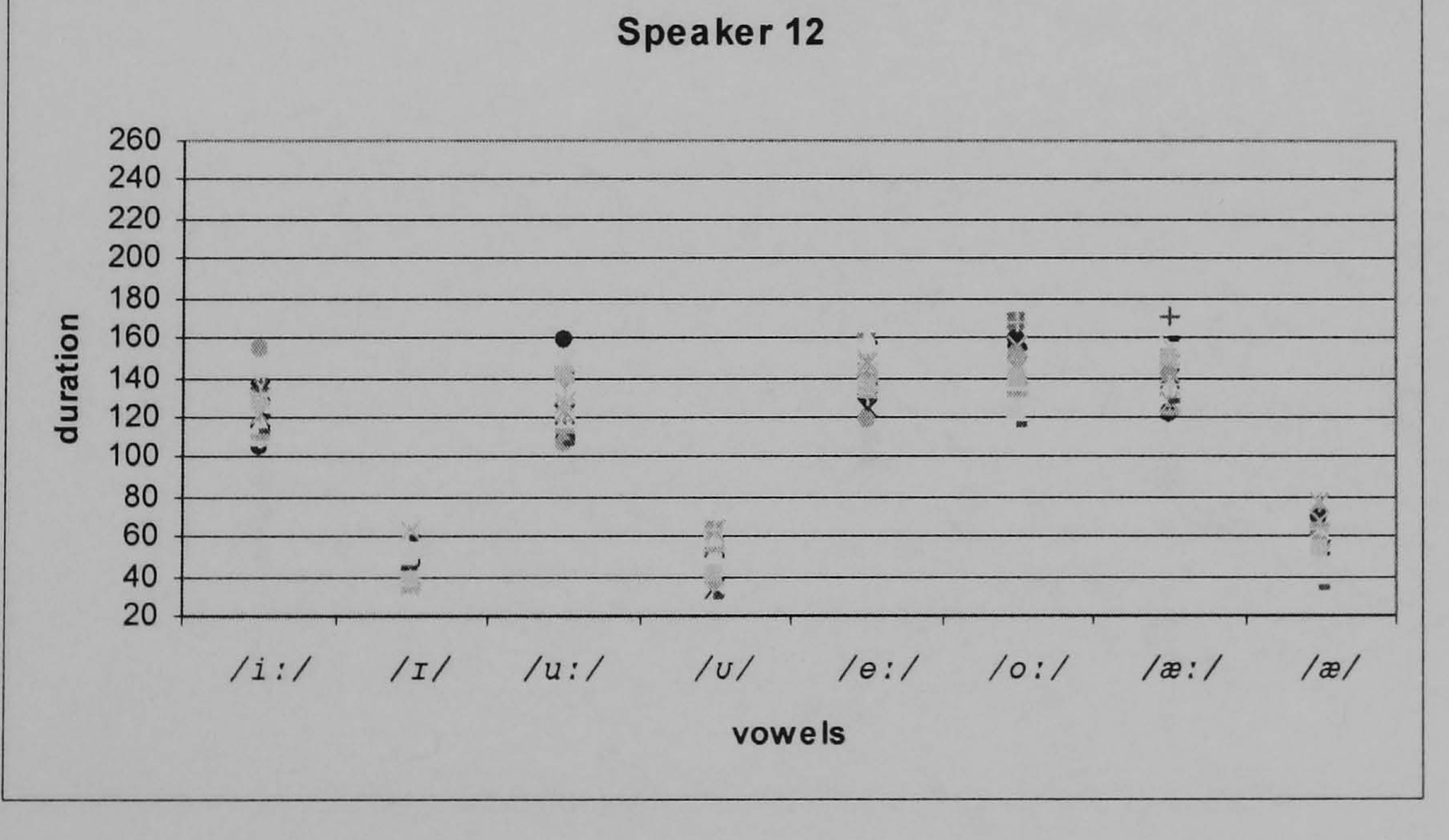

Speaker 7
260
240
220
200
180
160
140
120
100
80
60
40
20
duration
/i:/
/ɪ/
/u:/
/ʊ/
/e:/
/o:/
/æ:/
/æ/
vowels

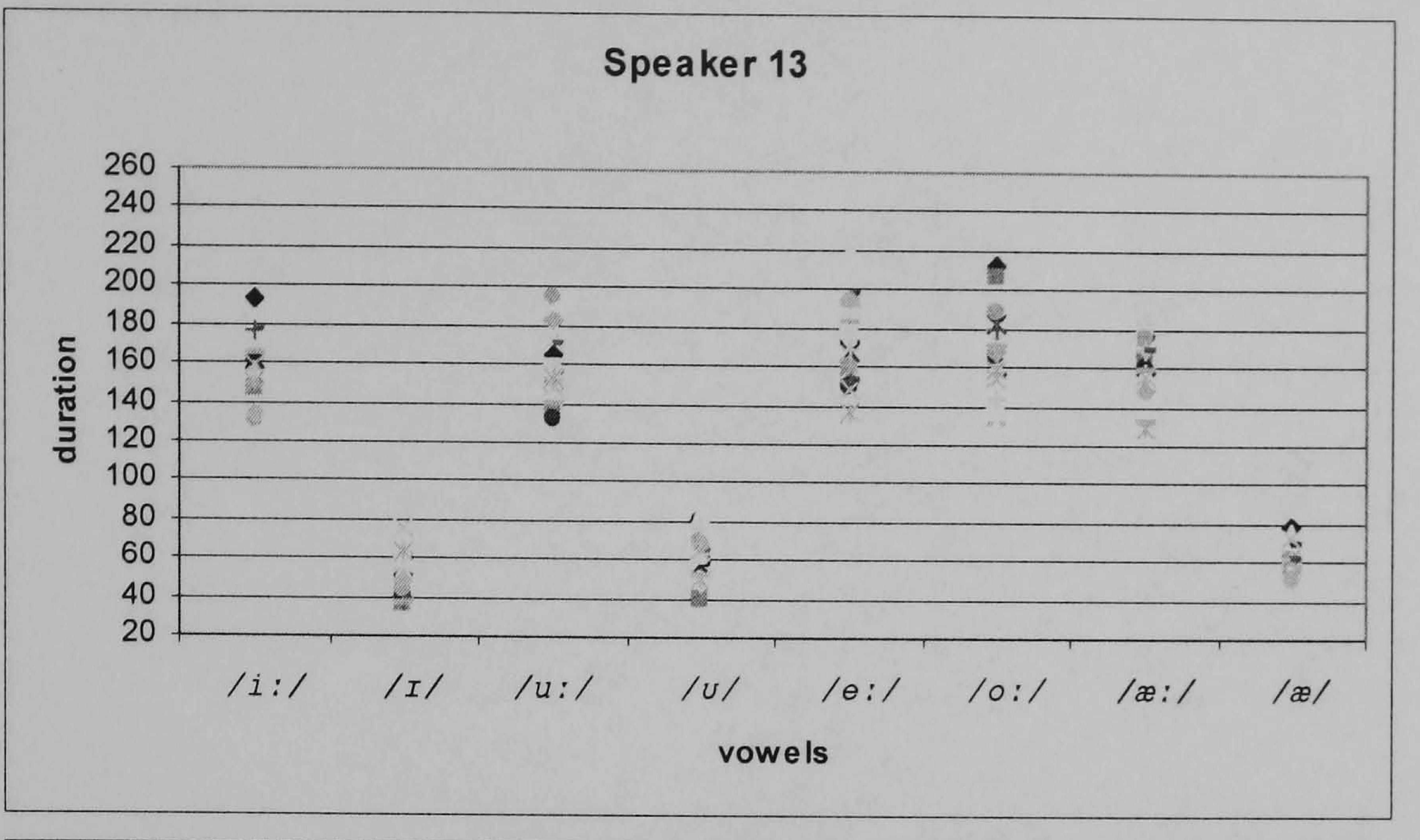

Speaker 8
260
240
220
200
180
160
140
120
100
80
60
40
20
duration
/i:/
/ɪ/
/u:/
/ʊ/
/e:/
/o:/
/æ:/
/æ/
vowels

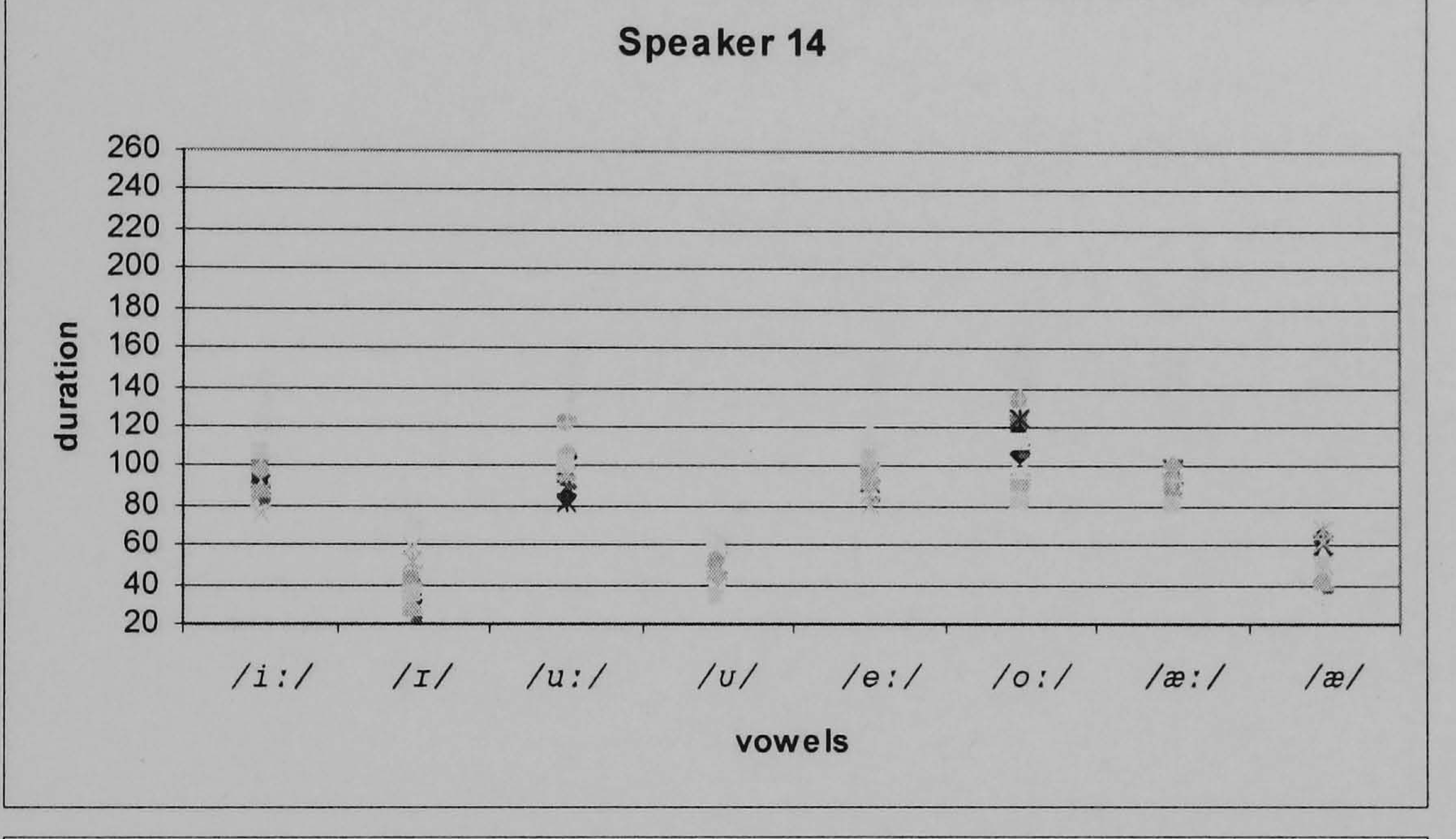

Speaker 9
260
240
220
200
180
160
140
120
100
80
60
40
20
duration
/i:/
/ɪ/
/u:/
/ʊ/
/e:/
/o:/
/æ:/
/æ/
vowels

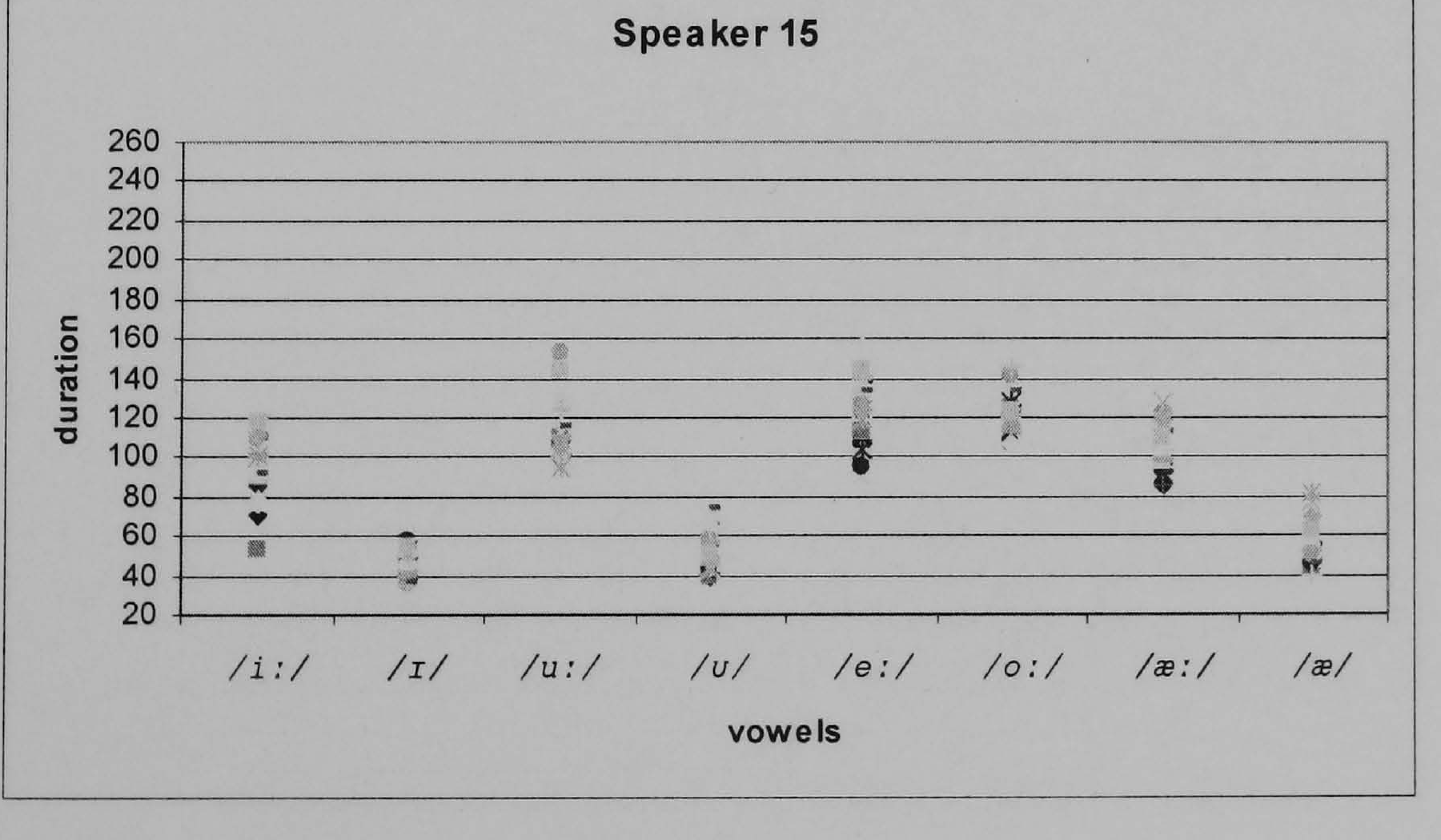

Speaker 10
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels

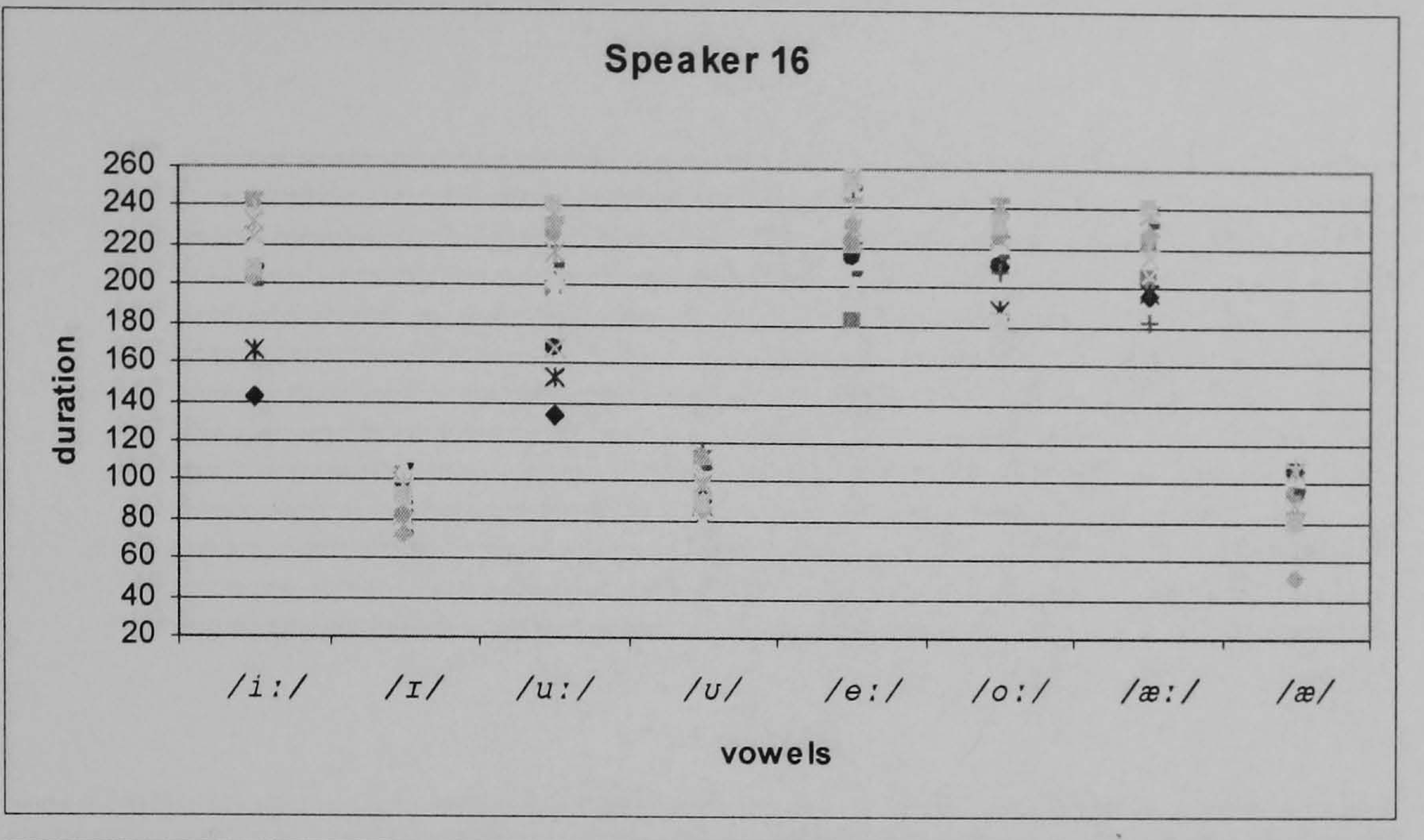

Speaker 11
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels

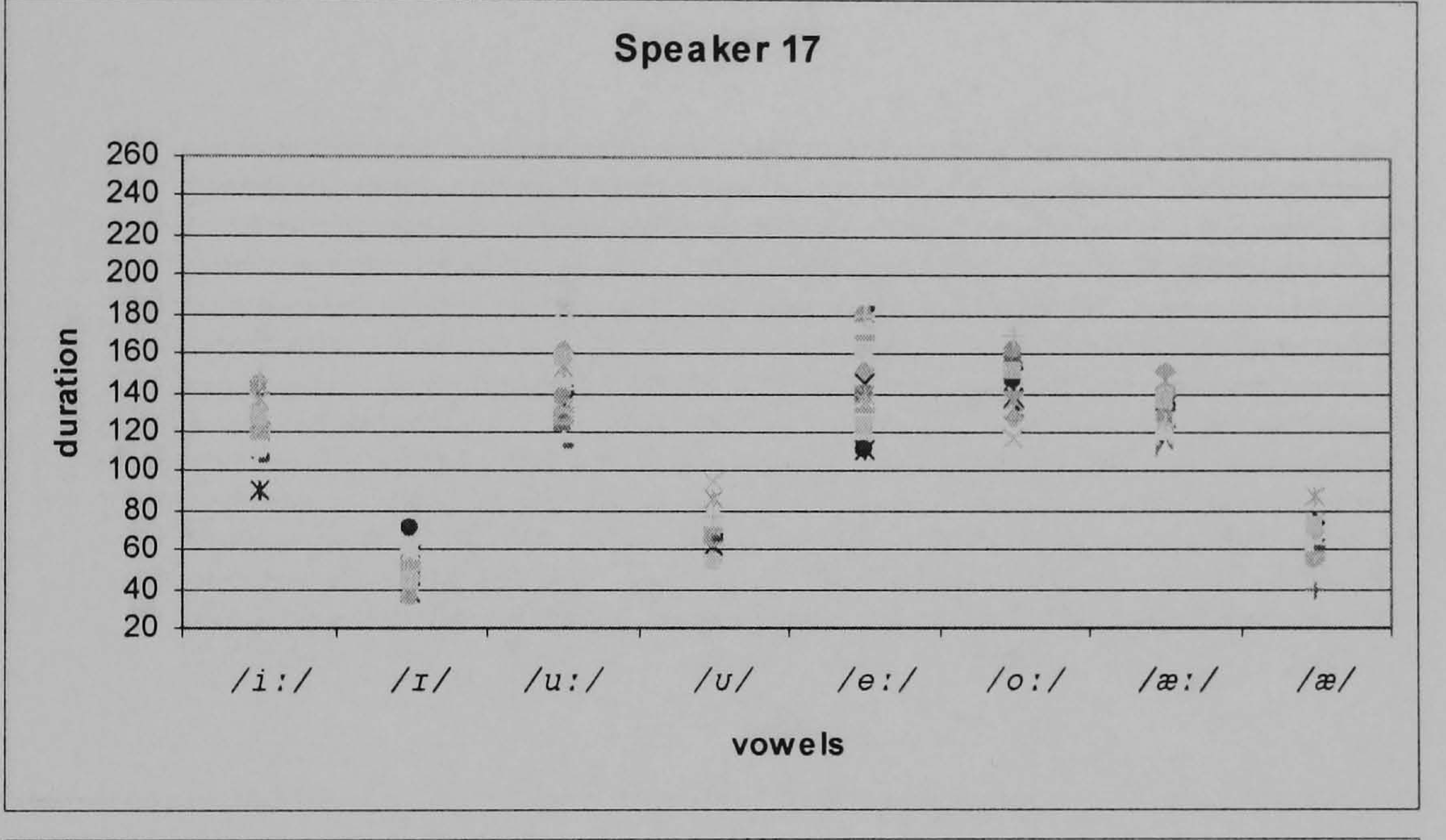

Speaker 12
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels

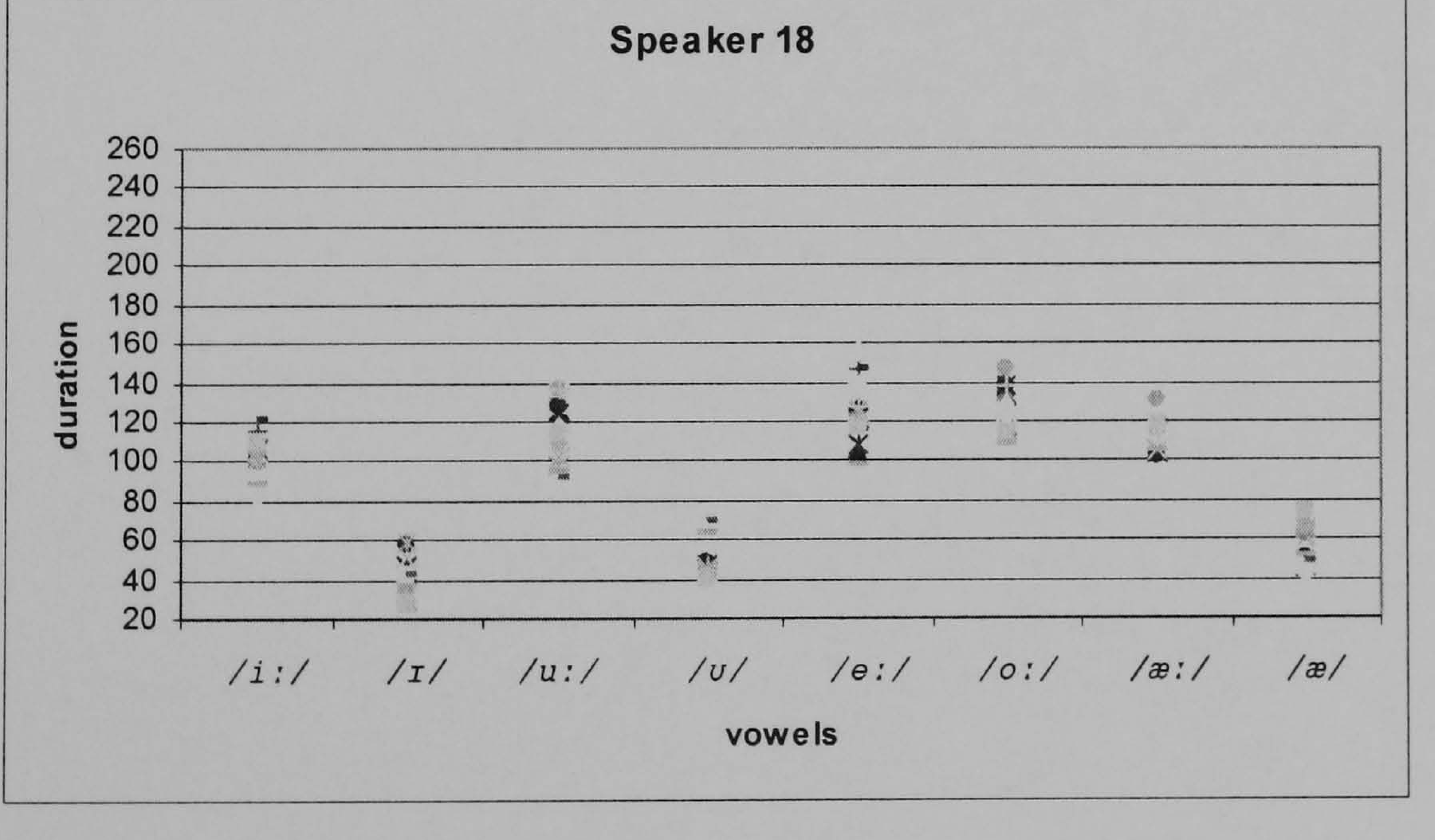

Speaker 13
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels

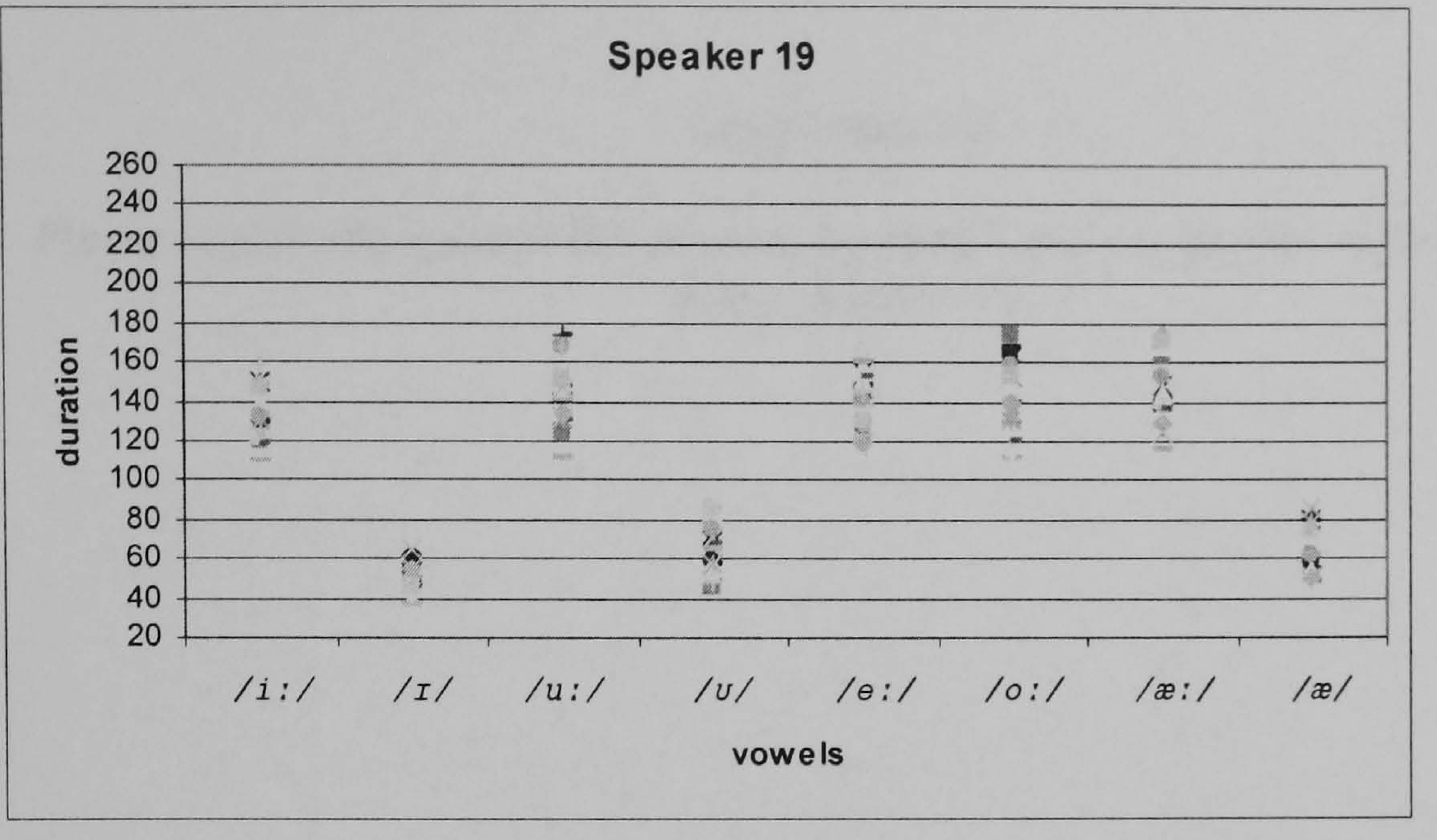

Speaker 14
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels

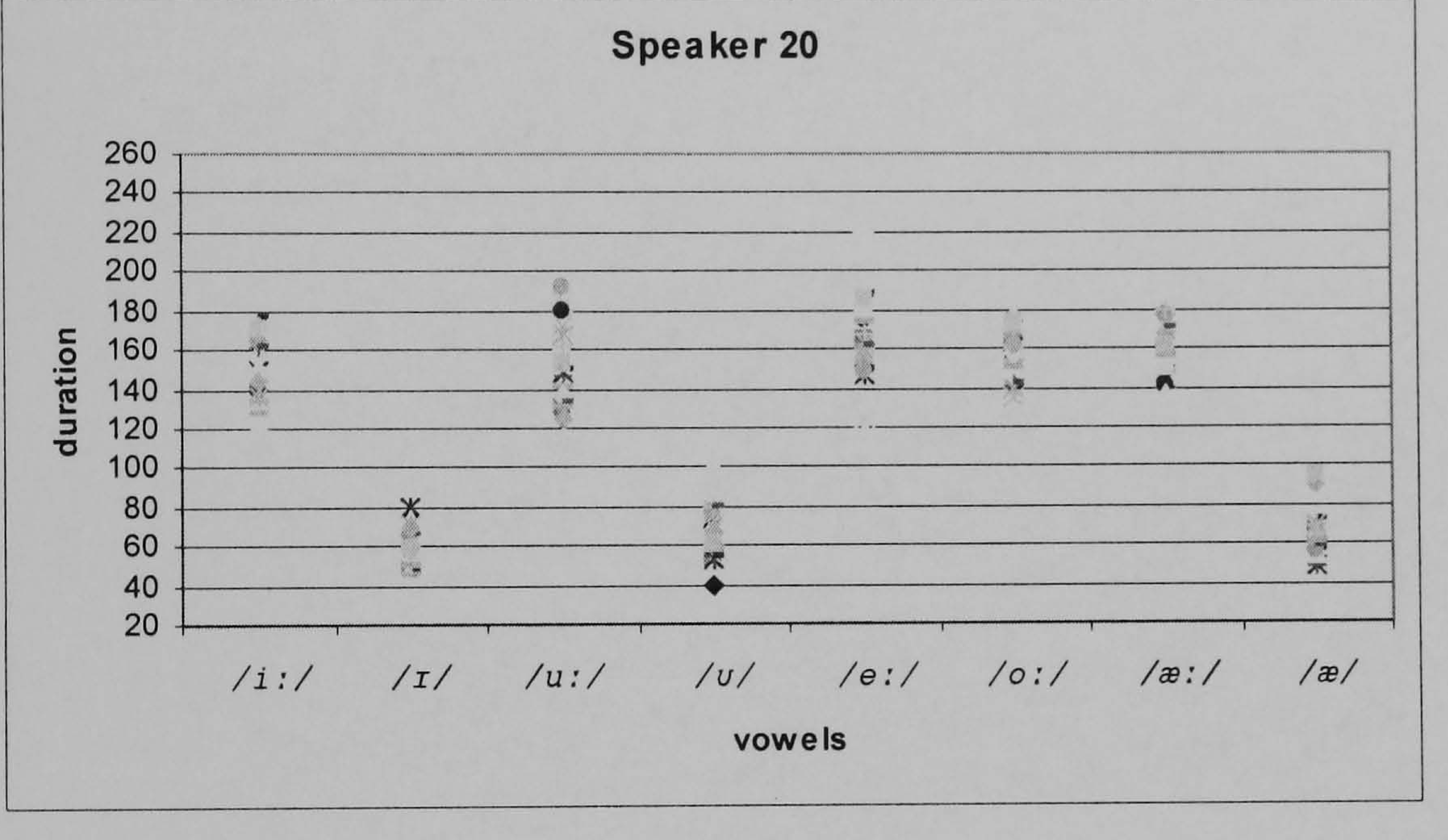

Speaker 15
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/ /ɪ/ /u:/ /ʊ/ /e:/ /o:/ /æ:/ /æ/
vowels


Speaker 16
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/
/ɪ/
/u:/
/ʊ/
/e:/
/o:/
/æ:/
/æ/
vowels


Speaker 17
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/
/ɪ/
/u:/
/ʊ/
/e:/
/o:/
/æ:/
/æ/
vowels


Speaker 18
duration
260
240
220
200
180
160
140
120
100
80
60
40
20
/i:/
/ɪ/
/u:/
/ʊ/
/e:/
/o:/
/æ:/
/æ/
vowels

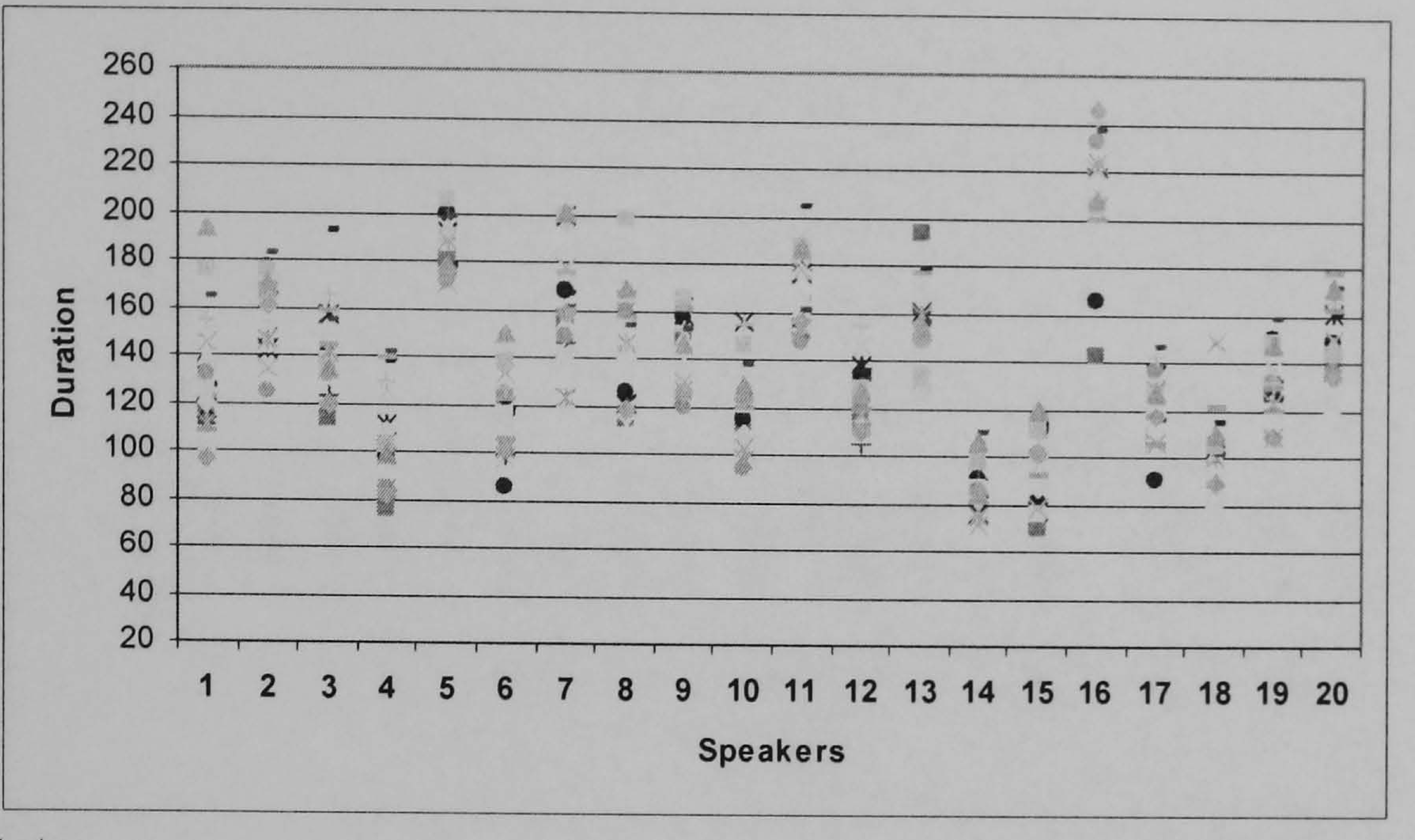

Speaker 19
260
240
220
200
180
160
140
120
100
80
60
40
20
duration
/i:/
/ɪ/
/u:/
/ʊ/
/e:/
/o:/
/æ:/
/æ/
vowels


Speaker 20
260
240
220
200
180
160
140
120
100
80
60
40
20
duration
/i:/
/ɪ/
/u:/
/ʊ/
/e:/
/o:/
/æ:/
/æ/
vowels

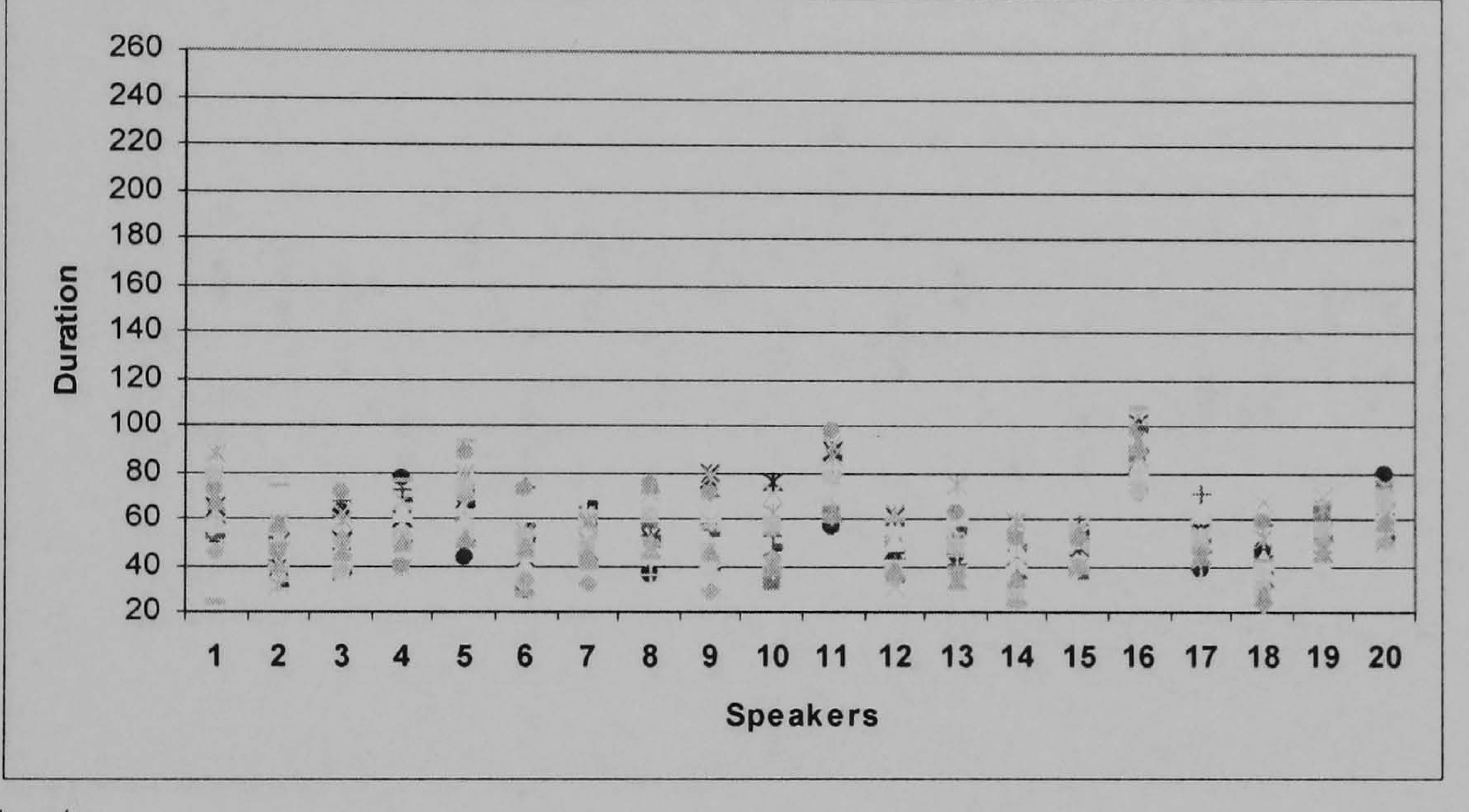
APPENDIX 3B

Planes representing vowel duration in ms of all vowels as produced by the whole group of speakers

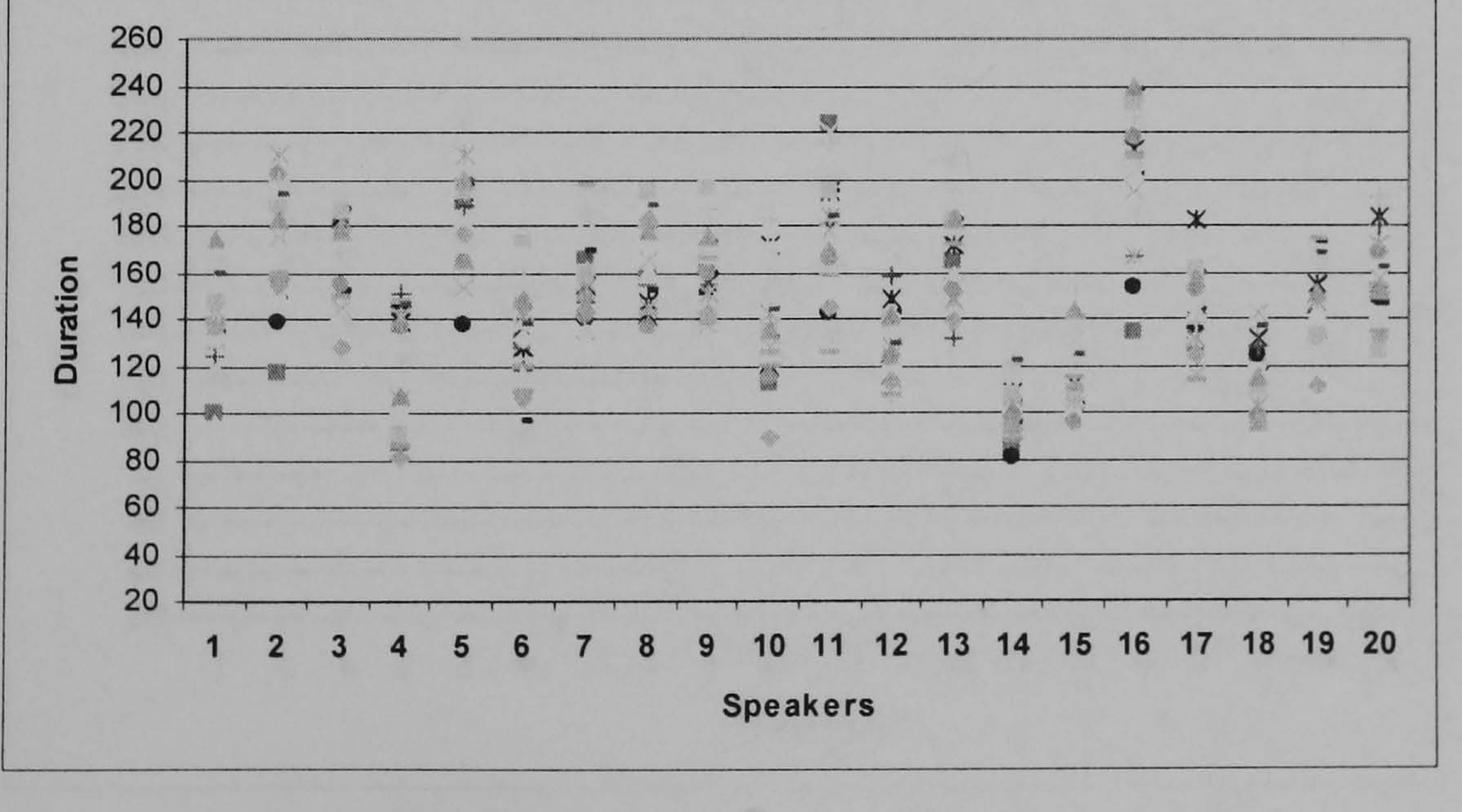
/iː/



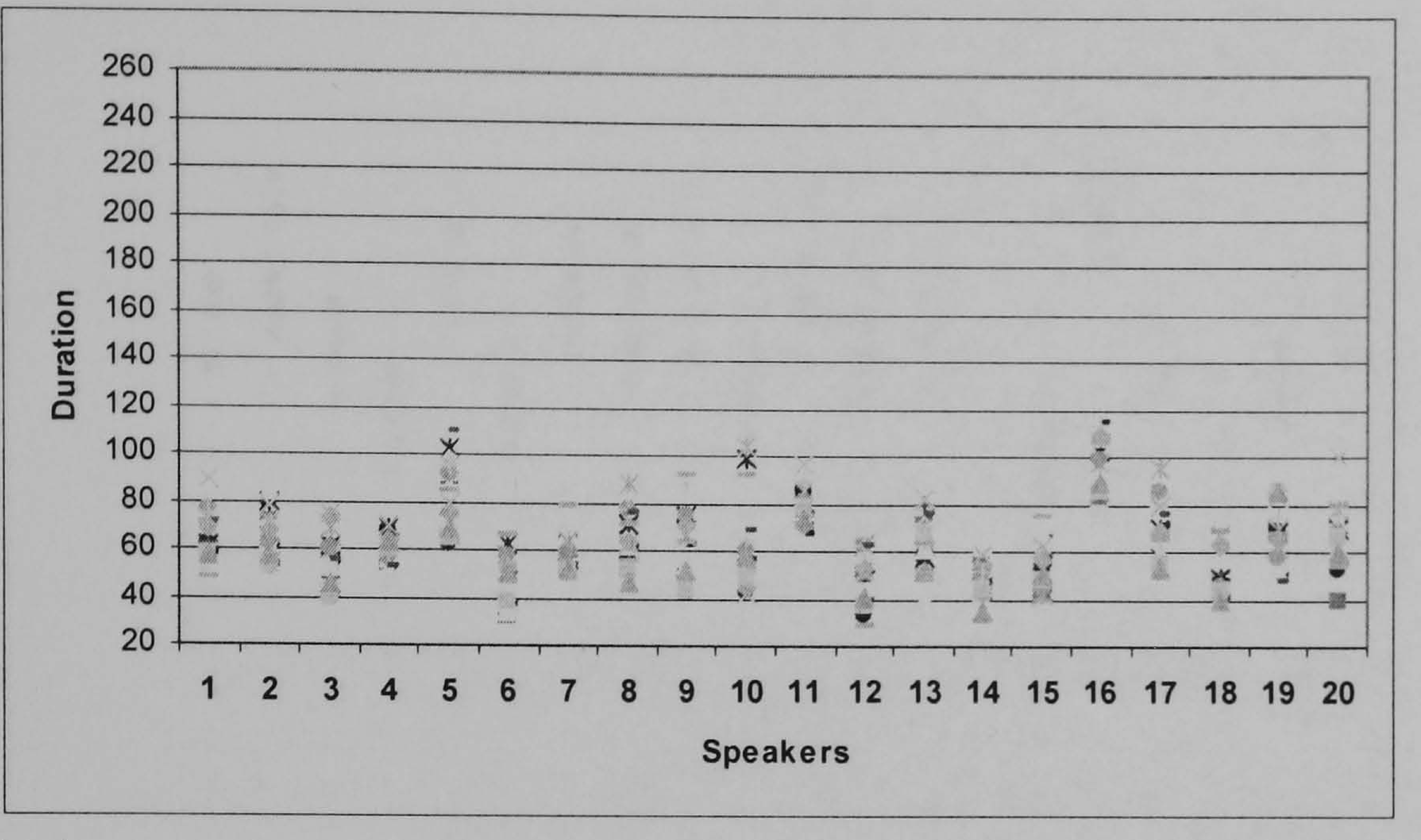
/ɪ/



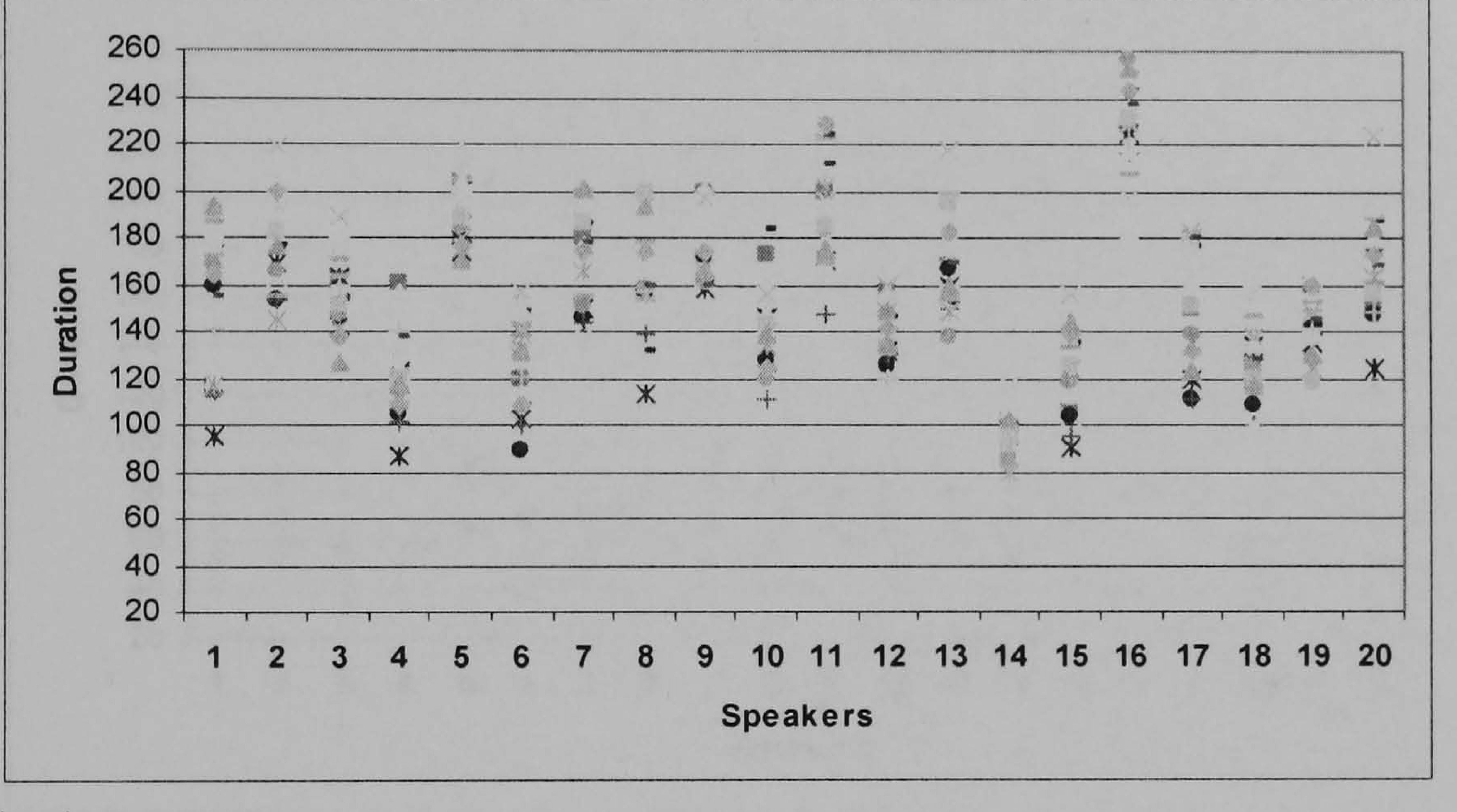
/uː/



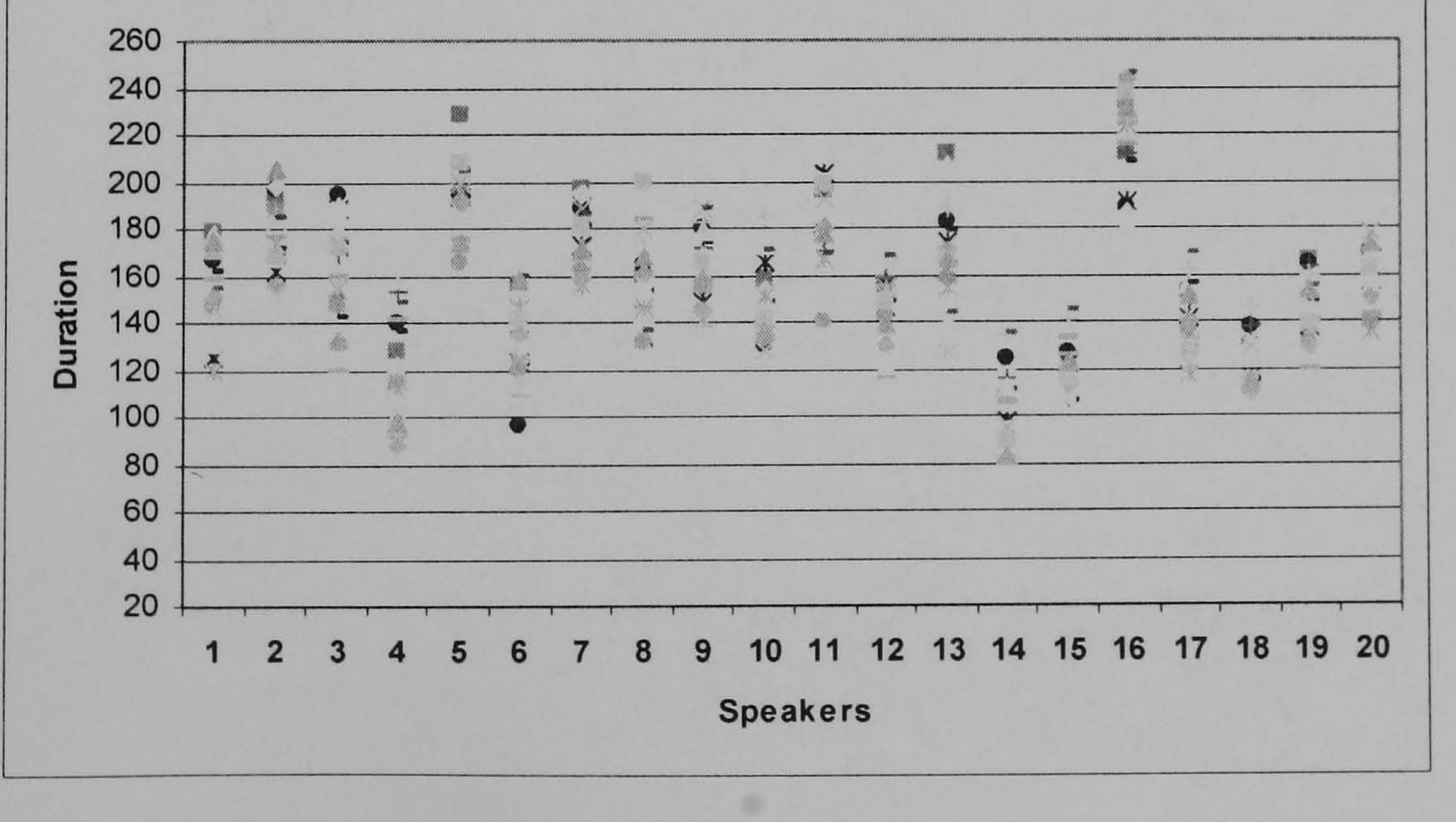
/ʊ/


Duration
260
240
220
200
180
160
140
120
100
80
60
40
20
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Speakers

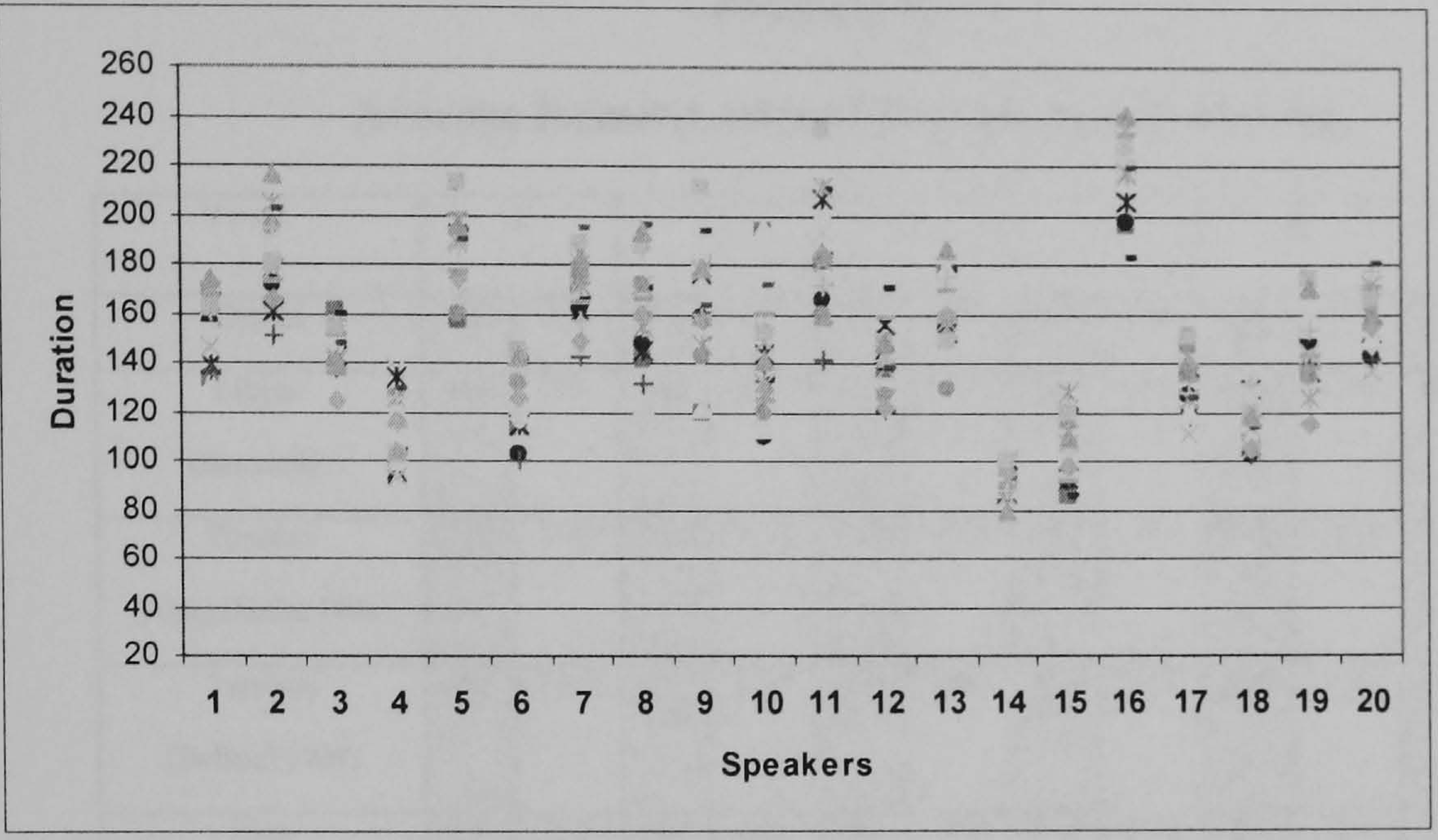
/eː/


Duration
260
240
220
200
180
160
140
120
100
80
60
40
20
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Speakers

/oː/

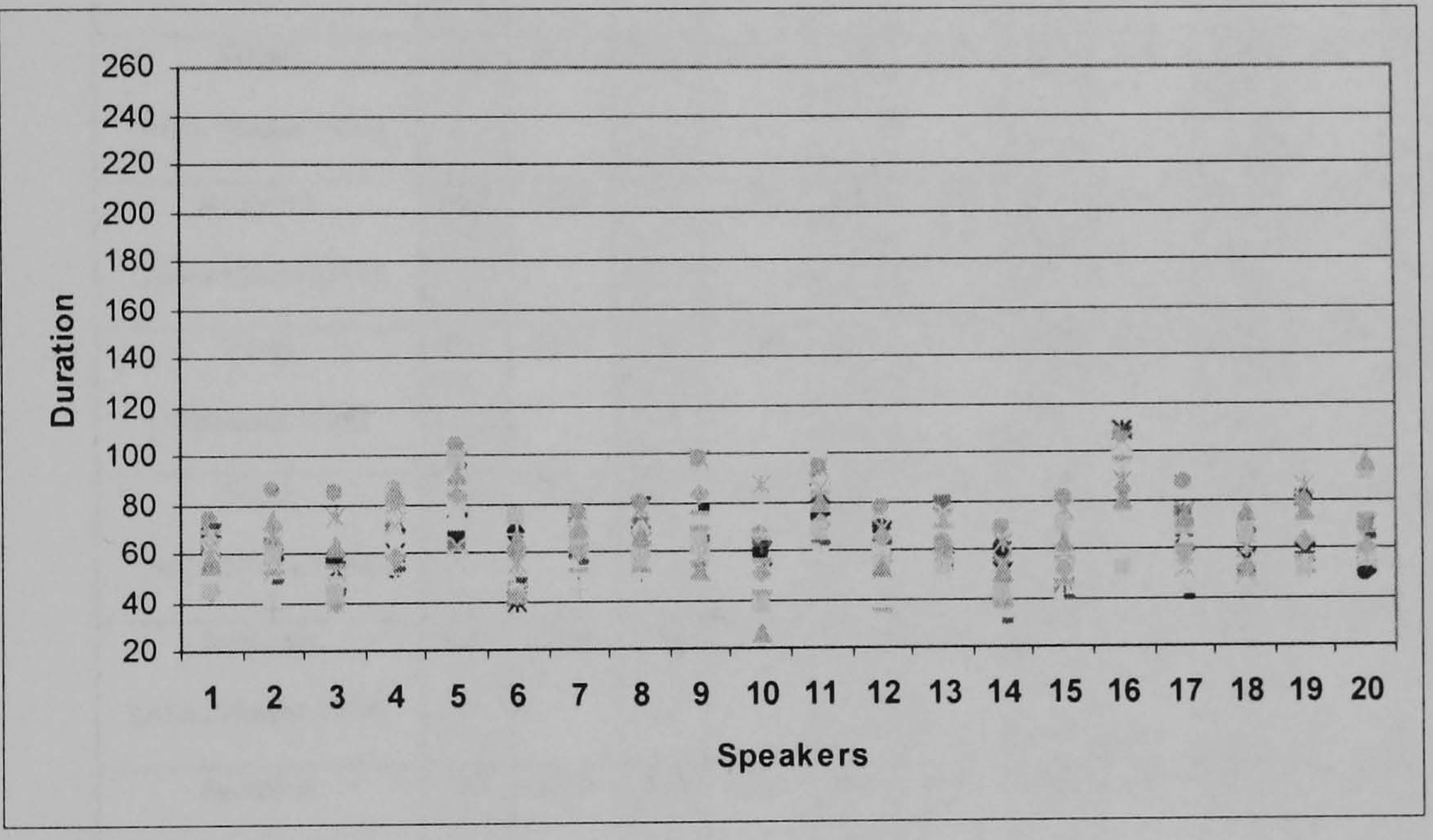

Duration
260
240
220
200
180
160
140
120
100
80
60
40
20
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Speakers

/æː/



260
240
220
200
180
160
140
120
100
80
60
40
20
Duration
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Speakers

/ə/



260
240
220
200
180
160
140
120
100
80
60
40
20
Duration
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Speakers

APPENDIX 4

Average formant values for some Arabic dialects

Vowel	i		i:		a		a:		u		u:	
Dialect	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
Libyan (this study)	404	1856	342	2214	555	1541	588	1641	443	1026	416	907
Tunisian (Abou Haidar 1994)	510	1690	315	2275	650	1590	610	1780	540	1135	360	830
Tunisian (Belkaid 1984)	355	1830	285	2195	400	1640	425	1720	340	995	310	790
Iraqi (Al-Ani 1970)	290	2200	285	2200	600	1500	675	1200	290	800	285	775
Syrian (Abou Haidar 1994)	415	2135	330	2465	700	1680	710	1560	430	1200	320	620
Jordanian (Abou Haidar1994)	565	1720	320	2295	780	1620	770	1521	580	1240	260	795
Saudi (Alghamdi 1998)	402	1841	292	2286	573	1537	655	1587	451	1302	350	958
Saudi (Abou Haidar1994)	540	1830	305	2530	695	1590	730	1540	540	1190	375	930
Sudanese (Abou Haidar 1994)	420	2000	325	2220	660	1600	730	1500	455	1040	380	900
Sudanese (Alghamdi 1998)	331	2066	272	2255	525	1564	635	1492	354	1308	319	984
Egyptian (Alghamdi 1998)	357	1749	256	2175	468	1505	462	1677	370	1285	319	942
Cairene (Newman 2002)	375	1575	290	1940	683	1435	610	1500	360	912	290	830
Qatari (Abou Haidar 1994)	500	1400	310	1990	620	1540	621	1280	490	1005	310	830
Lebanese (Abou Haidar 1994)	490	1530	280	2010	640	1390	610	1430	475	1060	330	795
Emirati (Abou Haidar 1994)	460	1720	335	2065	640	1660	730	1380	475	1075	350	990

APPENDIX 5
Auditory Analysis

Target	i:				ɪ				u:				ʊ				e:				o:				æ:				ə			
speaker	i:	ɪ	e:	other	ɪ	i	ə	other	u:	o:	ʊ/u	other	ʊ	u	o	other	e:	e	ɛ	other	o:	u:	o	other	æ:	a:	ɛ:	other	ə	æ	ɛ	other
1	x				x				x				x				x				x				x					x		
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x				x			
	x				x				x						x		x				x				x				x			
	x				x				x				x				x					x			x				x			
	x				x				x				x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x				x			
2	x				x				x				x				x				x				x				x			
	x					x				x			x				x				x				x				x			
	x					x			x				x				x				x				x				x			
	x					x			x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
3	x				x					x			x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x				x			

4	x				x					x			x				x				x				x				x			
	x					x			x				x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
5	x					x			x				x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x				x			
6	x				x				x				x				x				x				x				x			
	x				x				x					x						ai	x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x							ai	x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x							ai	x				x				x			
	x				x				x				x							ai	x				x				x			
	x				x				x				x							ai	x				x				x			
	x				x				x				x							ai	x				x				x			
	x				x				x				x				x				x				x				x			
7	x				x				x				x				x				x				x					x		
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x					x		

	x			x			x			x			x			x			x			x			
8	x			x			x			x			x			x			x				x		
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x				x		
	x			x			x			x			x			x			x				x		
	x			x			x			x			x			x			x				x		
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x				x		
	x			x			x			x			x			x			x				x		
	x			x			x			x			x			x			x			x			
9	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x				x		x			x			
	x			x			x			x			x			x			x				x		
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
10	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x			x			
11	x			x			x			x			x				x		x			x			
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x				x		
	x			x			x			x			x				x		x			x			
	x			x				x		x			x			x			x			x			
	x			x			x			x			x			x			x				x		
	x			x			x			x			x			x			x			x			
	x			x			x			x			x			x			x				x		
	x			x			x			x			x			x			x			x			

	x				x				x				x				x				x				x				x			
12	x				x				x				x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x						x		x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
13	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x				x			
14	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
15	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			

	x				x				x				x				x				x				x					x		
16	x				x				x					x			x				x				x				x			
	x				x				x				x				x					x			x					x		
	x				x				x				x				x				x				x					x		
	x				x				x					x			x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x					x			x				x				x				x			
	x				x				x					x			x				x				x				x			
	x				x				x					x			x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x					x			x				x				x					x		
17	x					x			x				x				x				x				x				x			
	x					x			x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
18	x				x				x				x							i:	x				x				x			
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x							ʌ
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
19	x				x				x					x			x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x					x				x			x				x				x					x		
	x				x				x					x			x				x				x				x			
	x				x					x			x				x				x				x				x			
	x				x				x							e	x				x				x				x			

	x				x				x							u:	x				x				x				x			
20	x					x			x				x				x				x				x				x			
	x				x				x				x				x				x				x					x		
	x				x				x				x				x				x				x				x			
	x				x				x				x				x						x		x						x	
	x				x				x				x							i:	x				x					x		
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x				x			
	x				x				x				x				x					x			x					x		
	x				x				x				x				x				x				x				x			
	x				x				x				x				x				x				x							a:
total	200				192	8			178	21	1		188	10		2	192			8	193	6	1		200				159	38	1	2
percentage	100%				96%				89%				94%				96%				96.5%				100%				79.5%			

Appendix 6

Glossary of the words recorded from participants in the production experiment

1. Words containing vowel /iː/

Entry	Word category	Meaning
/biːk/	Prep.+ pron.	by you
/diːn/	N.	religion
/giːs/	V. imp.	measure
/liːn/	N.	softness
/siːd/	N.	master
/siːn/	N.	an alphabetical letter
/tiːn/	N.	figs
/riːm/	N.	dear
/kiːs/	N.	bag
/liːg/	V. imp.	stop it

2. Words containing vowel /ɪ/

Entry	Word category	Meaning
/fɪn/	V. imp.	turn round oneself quickly
/bɪn/	N.	son of
/fɪz/	V. imp.	get up and go away
/hɪz/	V. imp.	shake
/kɪis/	V. imp.	said to a cat meaning go
/lɪm/	V. imp.	gather
/lɪz/	V. imp.	expel
/mɪd/	V. imp.	hand
/mɪs/	V. imp	touch
/tɪm/	V. imp	finish

3. **Words containing vowel /u:/**

Entry	Word category	Meaning
/bu:m/	N.	owls
/du:m/	V. imp.	be continuous
/lu:m/	V. imp.	blame
/mu:s/	N.	knife
/mu:t/	V. imp.	die
/ru:f/	V. imp.	be kind
/ru:m/	V. imp.	abide
/ru:z/	V. imp.	to hold something to guess its weight
/tu:t/	N.	mulberry
/bu:k/	N.+ pron.	your father

4. **Words containing vowel /ʊ/**

Entry	Word category	Meaning
/bʊk/	V. imp.	to cause something to explode
/bʊl/	N.	stamp
/fʊl/	N.	jasmine
/bʊd/	N.	avoidance, mainly used in the negative form /labud/ meaning unavoidably
/bʊg/	V. imp.	to stab with a knife, to open a can of a drink with a knife
/bʊm/	N.	sound made by the explosion of something
/bʊn/	N.	coffee
/dʊb/	N.	bear
/hʊm/	Pron.	they
/kʊb/	V. imp.	pour

5. **Words containing vowel /e : /**

Entry	Word category	Meaning
/be : n/	N.	Terribleness
/be : t/	N.	home
/se : f/	N.	sword
/se : r/	N.	strap
/de : n/	N.	debt
/de : r/	N.	doing
/ʒe : b/	N.	pocket
/ze : d/	N.	proper noun
/ze : n/	N.	beauty
/ke : l/	N.	weighing

6. **Words containing vowel /o : /**

Entry	Word category	Meaning
/do : r/	N.	level, wander
/lo : n/	N.	colour
/lo : z/	N.	almonds
/fo : z/	N.	wining
/ko : n/	N.	universe
/do : m/	N.	forever
/mo : t/	N.	death
/ro : z/	N.	the act of holding something in order to guess its weight
/ħo : ʃ/	N.	house
/go : s/	N.	bow, arch

7. Words containing vowel /a:/

Entry	Word category	Meaning
/læ:m/	V. past	blame
/sæ:d/	Adj.	enough
/fæ:t/	V. past	pass
/næ:b/	N.	canine tooth
/bæ:n/	V. past	appeare
/bæ:t/	V. past	stayed over the night
/sæ:m/	Adj.	poisonous
/dæ:m/	V. past	continue
/fæ:d/	Adj.	fed up of
/gæ:s/	V. past	measure

8. Words containing vowel /ə/

Entry	Word category	Meaning
/ləm/	V. past	gather
/səd/	V. past	(was) enough
/fəd/	V. past	fed up
/nəg/	V. past	nag
/səb/	V. past	curse
/fək/	V. past	take by force
/rən/	V. past	ring
/həm/	N.	worry, anxiety, distress
/məs/	V. past	touch
/wən/	V. past	(especially of electric devices) to make a continuous sound