INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.
2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.
4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

University Microfilms International
300 N. Zeeb Road
Ann Arbor, MI 48106

8409415

Rubin, Judith Ann

STATIC AND DYNAMIC INFORMATION IN VOWELS PRODUCED BY THE HEARING IMPAIRED

City University of New York Ph.D. 1984

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

STATIC AND DYNAMIC INFORMATION IN VOWELS PRODUCED BY THE HEARING IMPAIRED

by

JUDITH A. RUBIN

A dissertation submitted to the Graduate Faculty in Speech and Hearing Sciences in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

This manuscript has been read and accepted for the Graduate Faculty in Speech and Hearing Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

10-27-83
date

Katherine S Harris
Katherine S. Harris
Chairperson of Examining Committee

10-27-83
date

Irving Hochberg
Irving Hochberg
Executive Officer

Katherine S Harris
Katherine S. Harris

Harry Levitt
Harry Levitt

Nancy S. McGarr
Nancy S. McGarr

Michael Studdert-Kennedy
Michael Studdert-Kennedy
Supervisory Committee

Kenneth N. Stevens
Kenneth N. Stevens
Consultant

The City University of New York

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS

It is not unreasonable that doctoral candidates are given unlimited space for acknowledging the help of others in the completion of their thesis. We are all acutely aware that projects such as these simply cannot be done alone.

I would like to take this opportunity to thank my professors, past and present, who guided me in my education, encouraged my efforts and inspired me with their wisdom.

In particular, I extend my most sincere thanks to the members of my committee: Professor Harry Levitt, Dr. Nancy McGarr, and Professor Michael Studdert-Kennedy. I am deeply indebted to each and every one of them for their enthusiasm, suggestions, and guidance, and for their support throughout the course of this study.

A very special thanks must go to my advisor, Professor Katherine Harris, whose unending energy and interest, whose insightful comments and criticisms, and whose open-door policy, made the completion of this project one of the most fulfilling times in my academic career. I consider myself lucky to be counted as one of her students.

I would also like to thank:

Professor Kenneth Stevens, for graciously serving on my committee.

Margo Carter, for her patience and skill in preparing the figures for this document.

Gail Reynolds, for her editorial advice and assistance.

Russell Mills and D. H. Whalen, who generously offered their time and abilities with respect to writing the necessary computer programs.

Jean Sullivan, for the many hours she volunteered to help process data.

The staff and students at the Lexington School for the Deaf, St. Joseph's School for the Deaf, and St. Francis de Salles School for the Deaf. Without them this thesis could not have been completed.

Whenever one undertakes a project such as this, their must be a source of energy that is at all times positive and never once unsure. I would like to thank Steven Spitz for being all of that for me and much more. For his belief that I would finish when I was sure that I could not, for endless hours of listening and encouraging, for his sensitivity, advice, support and direction when I needed it most. . . I find it difficult to express my thanks but happily have a lifetime in which to try.

Judith A. Rubin

October 27, 1983

This work was supported in part by NIH Grants NS-17764, NS-13617, and HD-01994, and also by C. U. N. Y.

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The goal of many speech scientists is to discover the acoustic properties of the speech signal that make it uniquely intelligible to the human observer. In studying the speech of the hearing impaired we are more often faced with the question of why it is so uniquely unintelligible.

While straightforward listener transcriptions and ratings of the speech of the deaf have dominated the field , it has become clear that there are limitations to what we can learn from this experimental approach. As Monsen (1978) has noted, listener judgements can tell you how much of a child's speech is intelligible; they cannot tell you why.

More recently, acoustic and physiological measurements have been employed and provide us with information as to how articulatory activity in the speech of the deaf compares to that seen in the normal hearing talker. These studies may be categorized as focusing on:

1. static aspects of speech production (that is, fixed articulatory configurations or targets).
2. dynamic aspects of speech production (that is, movements from one

target to another, interarticulator timing, and changes in production as a function of phonetic environment).

The results of such investigations have led to several theories of vowel production in this population. Predictions, based on physical data, have been made as to how the measured differences in vowel production will affect the listener's ability to decode the acoustic signal and hence its overall intelligibility.

The purpose of this dissertation is to test systematically, both acoustically and perceptually, three hypotheses regarding vowel production in the hearing impaired.

Two of these hypotheses are concerned with issues related to target configurations. The first will be called the 'Absent Target Hypothesis' and suggests that hearing-impaired talkers do not have precise representations of articulatory/acoustic vowel targets. This supposedly results in their somewhat random selection of articulatory gestures for any given vowel. This hypothesis is based primarily on data averaged across talkers. Because of the apparent and excessive intersubject variability in this population, one must question the validity of theories based on such averaged data (Bush, 1981). Furthermore, it would be necessary to show a random association between vowel target and formant frequencies in order to substantiate a claim of an absence of vowel targets (Monsen, 1976a). This, of course, requires an accurate estimate of intrasubject variability: an aspect of deaf vowel production that has not been adequately assessed.

A second hypothesis related to vowel targets is based on individual data, and will be called the 'Deviant Phonology Hypothesis'. This hypothesis suggests that hearing- impaired speakers, as the result of limited auditory input, have abstracted deviant but well-defined articulatory targets. That is, they can be said to operate within a 'deviant phonological space' (Metz, 1980; Monsen, 1974, 1976d; Simon, 1983). This suggests that one will be able to find some feature(s) in the production of vowels which differentiates between targets and is produced with some degree of precision. As with the first hypothesis, validation would require measurements of variability in the vowel production of individual deaf speakers.

The notion of a deviant phonology implies that there is information systematically coded in the signal which may or may not be available to the uninformed listener. The perceptual implication is that a major cause of unintelligibility in the speech of the deaf is a mismatch between the phonology which guides the production of vowels by the hearing impaired and that assumed by the normal listener. One would predict, then, that correct identification of the speaker as hearing impaired, _a priori_ knowledge of the patterns of deviation in the speech of the deaf, and/or the opportunity to sample the novel phonological space by listening to a precursor sentence, would yield a significant perceptual advantage.

The Absent Target and Deviant Phonology Hypotheses represent two possible extremes. A logical alternative would be a target hypothesis which stands somewhere between the two. For example, hearing-impaired talkers may operate within the normal phonological system but with a reduced number of distinct targets. Many teachers of the deaf describe the vowel systems of their students in this way and may implicitly assume this in developing speech teaching strategies (McGarr et al., 1983). Nevertheless, this 'hybrid' hypothesis has not received direct attention in the literature.

A third hypothesis derives from acoustic and physiological measurements of the dynamics of vowel production in syllable context and focuses on time-varying rather than steady state aspects of production. This hypothesis will be called the 'Abnormal Coarticulation Hypothesis'. The extent and duration of formant transitions (Monsen, 1976b; Rothman, 1976), changes in articulatory gestures as a function of the phonetic environment (Boothroyd et al., 1974; Calvert, 1961; McGarr & Harris, 1983; Metz, 1980; Monsen, 1974; Nickerson, 1973; Osberger & Levitt, 1978; Whitehead & Jones, 1976) and measures of interarticulator timing (McGarr & Harris, 1983; McGarr & Löfqvist, 1982; Rothman, 1976) all suggest that hearing-impaired speakers do not co-produce segments in the same way as the normal hearing speaker nor (at least at the physiological level) in a way that is internally consistent. These findings have led to the notion that a major problem in the speech of the hearing impaired could be described as a 'breakdown in motor programming' (Harris & McGarr, 1980). While this is in sharp contrast to the notion of a stable but deviant phonology, the Abnormal Coarticulation Hypothesis and the Absent Target Hypothesis may not be mutually exclusive. Although they focus on different aspects of production, both suggest a certain degree of disorder and/or lack of structure in the system.

Monsen (1976d) has suggested that the lack of coarticulatory information in the speech of the deaf prevents the listener from decoding the signal in a normal way, thus rendering it uniquely unintelligible. The perceptual consequences of poor coarticulatory skills remains largely untested. The few studies which have attempted to relate the intelligibility of a given segment to measures of interarticulator timing (e.g. the timing of glottal opening with respect to oral articulation, the temporal relationship between orbicularis oris and genioglossus activity) have been unsuccessful (McGarr & Löfqvist, 1982; McGarr & Harris, 1983). Of course, the acoustic effects of interarticulator timing and their role in the intelligibility of normal speech is not well known. To test this third hypothesis, it would be more appropriate to choose a task in which the acoustic sources of coarticulatory information used by listeners in the perception of the test segment are qualitatively and quantitatively known.

One area where we do have a fair degree of established knowledge is normal vowel perception. At least three sources of articulatory/acoustic information may be relevant to the task: (1) steady state F1 and F2 frequencies (2) transitional information towards and away from the steady state formants and (3) information about the talker's vowel space and/or rate of speech. A systematic exploration of the listener's ability to use these three sources of information in decoding deaf vs. hearing vowels would be a direct test of the 'coarticulation' hypothesis. It would also serve as a forum for testing predictions as to the perceptual consequences of a possible deviant phonology.

The purpose of this thesis is to explore systematically the perception as well as the production of deaf vowels. The following specific questions were addressed:

1. How do hearing-impaired talkers compare to normal hearing controls regarding intrasubject variability in vowel production?
2. Does vowel production in the hearing impaired give evidence for a stable but deviant phonology or for the absence of articulatory/acoustic targets?
3. Do normal hearing listeners use the same acoustic sources of information in determining vowel identity in the speech of the hearing impaired as in the speech of normal hearing talkers?
4. Do experienced and inexperienced listeners differ in ability and/or strategy in identifying vowels in the speech of the deaf?
5. Do experienced and/or inexperienced listeners require dynamic vocal tract information in order to identify a speaker correctly as hearing impaired, and does identification of hearing status interact with phonetic decisions?

CHAPTER II

LITERATURE REVIEW

VOWEL PRODUCTION BY THE HEARING IMPAIRED

In an effort to determine the accuracy of vowel production in the congenitally severely and profoundly hearing-impaired population (hereafter referred to as either 'hearing-impaired' or 'deaf') both perceptual and physical measurement strategies have been employed. The results of this work have led to descriptions of error patterns in vowel production by the hearing impaired as well as theories regarding the underlying vocal tract gestures employed by these talkers in vowel production.

PERCEPTUAL MEASURES

Experienced and inexperienced listeners have been asked to transcribe vowels read in the context of simple sentences (Gold, 1978; Hudgins & Numbers, 1942; Smith, 1972), read in the context of mono- and polysyllabic words (Markides, 1970; Suonpaa & Aaltonen, 1981), read in a fixed /hVd/ context (Angelocci, Kopp & Holbrook, 1964; McGarr & Gelfer, 1983), and spontaneously produced in words (Gulian & Hinds, 1981; Nober, 1967).

In addition to their differences in test stimuli and elicitation procedure, these studies also differed in their criteria for a correct response, in the experience of their listeners, in the degree and type of hearing loss sustained by the children or adults serving as talkers, and in the method of speech training to which the talkers had been exposed. Thus, it is difficult to compare the absolute value of the vowel error rates reported.

Nevertheless, perceptual studies do report remarkably homogeneous error patterns. Levitt & Stromberg (1983) indicate that the most common error is a tense/lax substitution, followed by substitutions involving near neighbors (that is, vowels which are adjacent on the vowel triangle). Generally, substitutions are toward a more central vowel and back vowels tend to show less errors than front vowels except when the intended vowel is diphthongized.

Suonpaa & Aaltonen (1981) report different error patterns in the Finnish vowel system. In this study, /ɑ/, /u/, and /i/ had the lowest error rates of the test vowels. The rounded front vowels /y/ and /œ/ were most often misidentified. Suonpaa & Aaltonen (1981) analyzed vowel confusions according to place (front, harmonic front, back) and manner (close, half close/half open, open) of production. Contrary to earlier findings, it was the close front and back vowels (/i, u, y/) that were most often confused with one another (cross - triangle confusions). For example, labelling /u/ as /y/ was one of the most common errors made.

Based on such perceptual findings, Hudgins & Numbers (1942) suggested that vowel error patterns reflected inaccuracy or failure of the articulatory processes. A similar view was held by Smith (1972) who proposed that hearing-impaired children have a "poor appreciation of the quality of most vowels, never aim for a true vowel target, and naturally produce a neutral vowel as a substitute for every other vowel..." (p.106). That is, articulatory/acoustic targets should be considered absent. While intriguing, this hypothesis has not been adequately tested. It is clear that validation would require physical data since one cannot infer the details of articulation from percent correct perceptual scores.

The need for concurrent physical and perceptual measurements may be further illustrated by considering the perceptual requirements for vowel vs. consonant recognition. That is, the criterion for identifying vowels may be less strict than that for consonants (Hudgins & Numbers, 1942; Levitt et al., 1980). Thus, the frequently cited result that vowel identification is better than that for consonants does not necessarily lead to the conclusion that articulatory control for the production of vowels is superior to that for consonants. Furthermore, it has been pointed out that when a segment is perceived 'incorrectly' [1], it is impossible to know where in the articulatory gesture the presumed production error has occured (Levitt & Stromberg, 1983; Monsen, 1978). Finally, under certain conditions, those familiar with the speech of the hearing impaired (experienced listeners) achieve higher percent correct scores than naive listeners. Obviously, we do not want to conclude from this that the production skills of deaf talkers change from one listening condition to another. In short, issues regarding mediating perceptual processes must also be addressed if we are to understand how hearing-impaired talkers encode information in their articulatory gestures.

Only one study involving hearing-impaired talkers has looked at this issue by considering perceptual information obtained as a function of time (Calvert, 1961). Experienced listeners were exposed to six types of vocalic stimuli and asked, in each case, to identify the speech as produced by

1. a normal hearing talker
2. a hearing-impaired talker
3. a normal hearing talker imitating 'deaf speech'
4. a person with a 'voice disorder'

The six stimulus types were

1. an 8 word sentence ("I went to a party last Friday night.")
2. bisyllabic nonsense syllables (/həCVk/)
3. /hVd/ nonsense syllables
4. diphthongs
5. vowels produced in isolation
6. vowels gated out of their original context

Calvert found that a listener's ability to perform this task decreased from condition 1 to 6 with a sharp drop in performance for stimulus types 5 and 6. He suggested, therefore, that the distinguishing acoustic characteristics of 'deaf speech' were encoded in articulatory movement rather than in any steady state aspect of the signal per se. Calvert also noted that overall durational distortions, common in the speech of the deaf, may be a significant factor in identifying the hearing status of the speaker.

Calvert's presentation of the speech of the hearing impaired in a series of controlled perceptual conditions is unique in its attempt to reveal the underlying perceptual processes used in decoding their speech. This, of course, is the standard experimental procedure used to learn how information is 'encoded and decoded' in normal speech.

Calvert's findings are revealing because they focus on the possible perceptual importance of deviant movement in deaf speech. They are provocative as well, in that they assign very little perceptual importance to the target configuration and/or the steady state source characteristics of the speech of the hearing impaired. No attempt has been made to replicate this perceptual study. However, acoustic and physiological data have been collected to evaluate these two aspects of production, and we turn to these in the next section.

PHYSICAL MEASUREMENTS

The Absent Target Hypothesis:

Spectrographic and Linear Predictive Coding analysis have both been used to measure the steady state values of f0, and of F1, F2, and F3 center frequencies for deaf talkers (Angelocci et al., 1964; Calvert, 1961; Bush, 1981; Horwich, 1977; Levitt & Stromberg, 1972; Osberger et al., 1979; McGarr & Gelfer, 1983; Monsen, 1976a, 1976b, 1976c, 1978;

Stein, 1980; Suonpaa & Aaltonen, 1982). Test tokens have included vowels spoken in words within various sentences (Monsen, 1976a, 1976b, 1978) or within a fixed carrier phrase (Bush, 1981), in /hVd/ words (Angelocci et al., 1964; McGarr & Gelfer, 1983), in /həCVk/ nonsense syllables (Calvert, 1961) and vowels produced in isolation (Suonpaa & Aaltonen, 1982).

Physiological measures have also been used to examine steady state target configurations in the production of vowels by the hearing impaired. Such studies are, of course, the most direct way to learn about vocal tract gestures but are limited by the small number and age range of subjects who can be tested. Electromyographic recordings from hooked-wire and surface electrodes (McGarr & Harris, 1983; McGarr & Gelfer, 1983) and cinefluorography (Boone, 1966; Crouter, 1960; Stein, 1980) have been used to describe the positions, and coordinated movements of the lips, tongue, jaw, and larynx in vowel production.

Calvert (1961) and Angelocci et al. (1964) present their acoustic data averaged over talkers while the other studies report results for individual subjects. With the exception of Stein (1980) and McGarr & Gelfer (1983), acoustic analyses are restricted to mean values for talkers, the number of within-subject repetitions ranges from 1 to 5, and intrasubject variability is not discussed.

All the studies support the notion that articulatory movements for vowel differentiation do not follow the normal pattern. One commonly held view is that hearing-impaired talkers move their tongues little if at all for vowel differentiation (that is, the notion of 'articulatory immobility'). This view is based on both physiological and acoustic data.

Boone's (1966) cinefluorographic findings, suggesting that hearing-impaired speakers tend to retract their tongues toward the pharyngeal wall, led to the notion of a 'tongue backed position'. More recently, physiological results have supported the notion of tongue fronting. Using cinefluorographic techniques, Stein (1980) looked at tongue and jaw positions in the production of /i,ɑ,u/. She showed essentially no distinction in the position of the tongue for /i/ vs. /u/ with the tongue remaining in the front of the mouth for both vowels. Appropriate differentiation of tongue height was noted for high vs. low vowels. By referencing all measures to the mandible rather than the maxilla, Stein showed that, in general, variations in tongue height were a result of jaw rather than lingual control. That is, the deaf appeared to use their jaws extensively, if not exclusively, to differentiate /i/ and /u/ from /ɑ/.

McGarr & Gelfer (1983) used electromyography to examine genioglossus and orbicularis oris activity. Almost no differentiation could be found in genioglossus activity for the test vowels. In conjunction with acoustic measurements revealing an inappropriately high F2 for /u/, these findings support a pattern of 'tongue fronting' rather than backing.

In all studies, tongue movement was limited, supporting the general notion of a 'static tongue body posture' proposed by Stevens et al. (1983). These authors suggest that the hearing impaired may adopt an average articulatory posture which constrains the range of tongue movement and thus restricts F2 variation in vowel production. Limited formant variation was, in fact, noted by Angelocci et al. (1964) who found that for deaf talkers, mean values of F1 and F2 for the test vowels were restricted to the 'middle' of the normal vowel space. The ranges of these two formants around the vowel triangle were, therefore, considerably reduced. For normals, the ranges were 655 Hz and 1715 Hz for F1 and F2, respectively; for the hearing impaired, ranges for F1 and F2 were 330 Hz and 1148 Hz, respectively [2]. The acoustic finding of reduced range supports the perceptual impression of vowel neutralization. Moreover, Angelocci et al. (1964) plotted F1 and F2 values for all tokens of all vowels produced by all the hearing-impaired talkers in their study and showed extreme overlap between vowel categories. Angelocci et al. offer these findings in support of the notion that hearing-impaired speakers do not have precise representations of articulatory/acoustic vowel targets. This, of course, is the same conclusion that Smith arrived at from perceptual data. Thus, there is acoustic data to support Smith's 'no target' hypothesis of vowel production in this population.

Monsen (1976a) disputes this point and states that formant overlap is not sufficient evidence for claiming an 'absence of vowel targets'. He points out that one would have to show a near random association between formant frequencies and the intended vowel in order to substantiate this claim. This, of course, would require an accurate estimate of intrasubject variability in formant frequency. Token-to-token variability in vowel production by the hearing impaired has not been adequately explored.

Only two studies offer data on intrasubject variability for target configurations. Both show patterns of increased but not system-wide variability. Stein (1980) used 3 repetitions of the point vowels /i,ɑ,u/ for 5 hearing-impaired talkers. All talkers produced /ɑ/ with appropriate F1 and F2 values and with low intrasubject variability. That is, all talkers produced /ɑ/ in a well-controlled way and distinguished it from the other two vowels. The talkers differed, however, regarding variability and overlap in the production of /i/ and /u/.

F1 frequency for /u/ was well-controlled in talkers who showed an inappropriately high-frequency and highly variable second formant for this vowel. /i/ production was normal. Thus, while tongue position appeared to be a target parameter for /i/, it seems that 'jaw raised' may have been the only salient aspect of /u/ production for these talkers.

For another talker, both F1 and F2 frequency for /u/ were well-controlled (low variability). The second formant frequency for /u/ was high and completely overlapping with that for /i/. Here, one might conclude that the articulatory targets for /i/ and /u/ were precisely the same. For this talker, it could be that jaw height was the only articulatory variable used in vowel differentiation and that a static tongue fronted position had been assumed. An analysis of intermediate vowels would be informative here.

A third talker's acoustic pattern of point vowel articulation showed the 3 vowels occupying mutually exclusive areas while being deviant with respect to the normal vowel space. This might suggest well-defined (though deviant) targets based on F1 and F2 frequency. The F1 and F2 ranges for this talker were considerably smaller than for the others mentioned. It would be inaccurate, however, to conclude that this talker had fewer vowel targets. In fact, based on the data presented, quite the opposite appears to be true.

The second study of target configurations for individual tokens (McGarr & Gelfer, 1983) obtained 10 repetitions of 6 test vowels for a single talker. In general, F2 range was restricted. The results also showed clear differentiation between /ɑ/ and the high vowels /i/ and /u/. As in Stein's study (1980), F2 values for /u/ were inappropriately high, overlapping with /i/ and highly variable. Both these studies show excessive token-to-token variability in the production of some but not all vowels. This suggests that while the relation between formant frequency and vowel target may be ambiguous, it is not random. These findings do not support the notion of an absence of vowel targets, but rather suggest a reduction in the number of distinct vowel targets. Intrasubject variability must be assessed in testing the validity of the Absent Target Hypothesis.

The Deviant Phonology Hypothesis:

In contrast to the Absent Target Hypothesis, it may be that the hearing-impaired speaker defines vowel targets in terms of articulatory/acoustic variables not normally used to specify vowel identity (that is, substitutes one parameter for another).

Monsen (1974, 1976a, 1976d) used individual data to argue for the extreme form of a substitution routine. In contrast to the Absent Target Hypothesis, he suggested that some hearing-impaired talkers, as a result of limited auditory input, may operate within a deviant but well-defined phonological space. Vowel targets may be based on alternative and more salient aspects of the signal (e.g. F1 alone, duration, nasality), but are clearly defined and their production well-controlled nonetheless. The emphasis on control in production is what distinguishes this hypothesis from others.

In Monsen's work (1976a), only the extremes of the normal vowel triangle were compared to determine formant frequency ranges. /i,ɑ,ɔ/ served as test vowels since they approximate the highest and lowest values for the two resonances in normal hearing speakers. In this case, then, the definition of range involves an a priori assumption regarding which vowels will represent the extremes of a given talker's vowel space.

Five tokens of each vowel were measured. For all but a few sample cases, only the mean values of these tokens were presented. Monsen describes a tendency toward immobility in F2; second formant values 'float' around 1800 Hz for many talkers. Sample individual F1/F2 plots show a 'reduced phonological space'. That is, a restricted range of formant values with overlap between vowel categories. Reduced range is, more often than not, due to second formant frequency, as is the overlap between vowel spaces. Therefore, not only does F2 vary little between one vowel target and another, but it is poorly controlled as well. However, F1 ranges for some talkers were of normal size and appeared to be vowel related (The opposite trend, well-differentiated F2's and poorly differentiated F1's, was never seen). Therefore, the vowel spaces may have been deviant in that they were defined by F1 frequency alone.

In a study of the tense/lax distinction, Monsen (1974) demonstrated that the i/I distinction was exaggerated with regards to the durational contrast and suggests that this feature of production might also be a substitute for articulatory changes.

Other studies have also supported the notion of a deviant phonology. Based on data averaged across talkers, Angelocci et al. (1964) showed that the range of f0 variability across vowel targets was greater for the deaf than for the normal hearing talkers. Taking this fact in conjunction with their findings of reduced formant range, the authors suggested that their talkers may have been substituting large variations in fundamental frequency for upper articulator movemement in an attempt to differentiate between vowels. The notion of substituting changes in fundamental frequency for articulatory gestures was also suggested by Horwich (1977) and by Nickerson (1975).

One cannot, however, adequately test the Deviant Phonology Hypothesis if acoustic data are pooled across talkers. That is, once data has been averaged, we cannot be sure that mean increases in one parameter (e.g. f0 range) and mean decreases in another (e.g. formant range) are due to the same talkers. Bush (1981) highlighted the problems that occur as a result of averaging across hearing-impaired subjects. She used 20 hearing-impaired talkers, 10 vowels, and 4 repetitions of each vowel (each repetition in a different phonetic context). In comparison with a normal hearing group of talkers, the deaf showed a greater f0 range around the vowel triangle (48Hz vs. 15 Hz), a slightly larger F1 range (405 Hz vs. 388 Hz), and a reduced F2 range (633 Hz vs. 1373 Hz) [3]. These data alone might have suggested the presence of a deviant phonology. Bush noted, however, a large amount of individual variability in the hearing- impaired group on these measures. Some talkers had normal F1 and F2 ranges, some had a greater than normal F1 range and reduced F2 movement, and some had reduced ranges for both measures. By grouping subjects appropriately, Bush showed that vowel related changes in f0 were positively, rather than negatively correlated with upper articulator movements. That is, those talkers who showed a greater range of F1 and F2 frequencies around the vowel triangle were also the talkers who demonstrated the largest vowel related changes in f0.

Therefore, while mean values are suggestive, to test the hypothesis of a deviant phonology would require a large number of intrasubject repetitions and individual rather than mean data. One would have to show two things:

1. that the token-to-token variability in the 'novel' parameter was significantly smaller than in the one it was replacing, and, perhaps, significantly smaller than (or at least equal to) the variability of that parameter in normal vowel production.

2. that the novel parameter served to distinguish between vowel targets in a way the ordinary parameter had not.

Imagine, for example, that a speaker had a monotone voice. Here, f0 variability would be small but its differentiating power negligible. If, on the other hand, a talker showed a large f0 range (highest - lowest mean value) but considerable token-to-token variability for most vowels, it would be difficult to argue that the talker based his vowel targets on this parameter. It is clear that the issue of intrasubject variability and the analysis of individual talkers is crucial to the validation of both the Deviant Phonology Hypothesis and the Absent Target Hypothesis.

The perceptual implication of a deviant phonological system is that information systematically coded in the signal will be unavailable to the naive listener familiar only with normal sound patterns. But what if listeners have the opportunity to sample the phonological space by hearing a precursor sentence? One would predict that this would improve their ability to interpret the signal. What of the listener who has had years of experience in decoding the speech of the deaf? We would predict that this individual would enjoy a considerable and measurable perceptual advantage over an inexperienced listener. The notion of an experienced listener advantage in the perception of deaf speech is therefore of considerable interest, and will be discussed below.

The Abnormal Coarticulation Hypothesis:

An alternative focus with respect to vowel production in the hearing impaired is the consideration of how these segments are produced in context. Calvert's (1961) findings suggested that this aspect of production was particularly revealing with regards to the peculiarities of deaf speech. In what ways do the dynamic vocal tract gestures of this population differ from the norm and of what relevance is this to the intelligibility of their utterances?

Second formant transition duration and extent were measured in the hearing impaired by Monsen (1976b) and by Rothman (1976). Both studies showed that the frequency extent of transitions tended to be reduced in comparison to that seen in hearing talkers. For the syllables /di/ and /bi/ (Monsen, 1976b), for example, the average F2 frequency change for deaf subjects were 160 HZ and 320 Hz, respectively. Comparative values in normal hearing subjects were 380 Hz and 705 Hz, respectively. While this can be explained, in part, by lower than normal F2 target frequencies for the vowels, Monsen shows examples of different transition patterns between the two groups for equal F2 targets.

Rothman (1976) reports slower rates of formant movement for the hearing impaired than for normal hearing talkers. Furthermore, he showed a reduction of the anticipatory and perseverative coarticulatory effects that were seen in normals. He states that hearing-impaired speakers tended to begin and end articulatory sequences in the same way regardless of the phonetic context; a phenomenon he calls 'articulatory stereotyping'.

Spectrographic measures have also been used to examine another presumed instance of coarticulatory behavior: duration as a function the phonetic and linguistic context (Boothroyd et al., 1974; Calvert, 1961; Metz, 1980; McGarr & Harris, 1983; Monsen, 1974; Nickerson, 1973; Osberger & Levitt, 1978; Reilly, 1979; Whitehead & Jones, 1976, 1978). A number of these studies showed evidence that hearing-impaired talkers do not vary vowel duration as a function of context as do normal hearing speakers. Others have suggested that a general trend towards normal variation may be evident but that the effects are quantitatively diminished and highly variable from subject to subject (Reilly, 1979; Whitehead & Jones, 1976, 1978). Whitehead & Jones grouped their subjects according to hearing loss and showed significant durational context effects for the less profoundly hearing-impaired group. Standard deviations for the three test groups (normal hearing, hearing-impaired, and deaf) were reported and reveal greater intersubject variability for both hearing-impaired groups in comparison to the normal hearing talkers. Reilly (1979) showed similar context effects for the normal hearing and hearing-impaired groups. She did, however, note that the standard deviations for duration measures were considerably larger for hearing-impaired than for normal hearing talkers. Therefore, due to large intersubject variability, averaging across hearing-impaired talkers may obscure actual effects.

A further aspect of coarticulation in speech production is that of interarticulator timing. Physiological studies of interarticulator timing in consonant-vowel sequences examine movements coordinated in time. A 'time-locked' nature to sequential gestures has been found for certain syllables in normal hearing speakers, with the relative timing of lip and tongue gestures being invariant across variations in stress and rate (Tuller & Harris, 1980; Tuller et al., in press). Hearing-impaired talkers, on the other hand, show a 'failure of interarticulator programming' (McGarr & Harris, 1983; McGarr & Löfqvist, 1982; Rothman, 1976). That is, the timing between lip and tongue activity, or between laryngeal and upper articulatory movement varies from token to token.

This suggests an inability to organize sequential gestures and a breakdown in motor programming. Of perhaps paramount importance here, are the findings of exceedingly high intrasubject variability at the physiological level. That is, talkers do not appear to substitute a stable but unique timing pattern for the norm. At least in this aspect of production, no support can be found for the notion of a stable but deviant phonology (Harris & McGarr, 1980).

Physiological and acoustic evidence suggest, then, that hearing-impaired talkers do not coproduce or coarticulate segments as does the normal hearing speaker. This aspect of production may best describe the underlying deviant nature of speech production by the hearing impaired. As will be discussed below, there is evidence to suggest that the normal listener perceives speech by continuously integrating information over time. Therefore, it may be that the absence or reduction of coarticulatory information in the acoustic signal prevents the listener from decoding the speech of the deaf in a normal way. We must now come full circle and say that it is surprising that we have no perceptual data to substantiate and/or elaborate on this hypothesis. If, in fact, the speech signal produced by these talkers is devoid of or has only minimal coarticulatory cues, one should be able to demonstrate this perceptually. The prediction would be that listeners will benefit less from being provided with a segment's surrounding articulatory/acoustic information than they do in the perception of 'normal speech'. Furthermore, those experienced with the speech of the hearing impaired may have learned to make better use of diminished coarticulatory cues and/or to ignore misleading information of this kind. Again then, the details of the experienced listener advantage are of interest.

LISTENER EXPERIENCE

Studies have shown that experience with the speech of the hearing impaired can be a significant factor in the results of perceptual measurements (Mangan, 1961; Markides, 1970; McGarr & Gelfer, 1983; McGarr, 1978; Nickerson, 1973; Thomas, 1963).

Mangan (1961) used monosyllabic words as stimuli and allowed the listeners (teachers, mothers, and naive subjects) to hear and see the talkers. She found a 14% difference between teachers and the other two groups. Thomas (1963) used both words and sentences and three groups of listeners (experienced, semi-naive, and naive). In general, the differences between listeners were greater for word than for sentence stimuli. The differences between experienced and semi-naive and between semi-naive and naive subjects in the perception of sentences were 2.9% and 7.4%, respectively. The equivalent comparisons for words were 7.3% and 19.2%, respectively.

Markides (1970) used spontaneous speech samples and scored experienced vs. inexperienced listeners on the percentage of total words spoken that were correctly understood. He found a 12% difference for deaf children (31% vs. 19% correct) and a 7% difference for hard-of-hearing children (83% vs. 76% correct). Nickerson (1973) showed significant differences between listeners for both read and spontaneous speech. He suggested that the differences were stable across a wide range of talker intelligibility.

McGarr (1978) used a more systematic approach in trying to uncover the nature of the experienced listener advantage. Test stimuli were high and low context sentences, words produced in isolation and words segmented from sentences. The overall effects of listener experience, as well as the interaction between groups of listeners and the above stimulus parameters were evaluated. For all conditions and types of stimuli, experienced listeners scored higher than inexperienced listeners. The differences were significant for all but the segmented word condition. Sentences showed the largest listener effect. The improvement in scores as a function of high vs. low context sentences and the _a priori_ intelligibility of the test words were equivalent in the two groups of listeners. This suggests that the listener advantage does not simply rely on the experienced listener's ability to make better use of linguistic contextual information. McGarr showed no consistent difference between the two groups of listeners in terms of relative variability, except when scores approached 0%. Since all listeners were unfamiliar with the children in the study, familiarity with the talker could not explain the perceptual advantage.

Boothroyd (1977) has shown that by using a forced-choice paradigm, the listener advantage can be reduced if not eliminated. That is, he shows significant listener effects for open set materials (% recognition of phonemes in C-N-C words, and % recognition of key words in sentences). With the same set of talkers and listeners, only a small listener advantage is seen in a Feature Identification Test using a two alternative forced-choice format. This testing paradigm effectively eliminates the influence of response criteria. However, since it is the details of experienced listeners' response criteria that we hope to discover, a forced-choice paradigm is uninformative.

Monsen (1978) asked experienced and naive listeners to listen to a series of four tapes and assessed the intelligibility of words produced in sentences. While the mean differences between listeners was 9%, a strong learning effect was apparent. Monsen showed an experienced listener advantage of 14% on the first test tape but an advantage of only 5% on the last. He claims that while experience with the speech of the hearing impaired "is a considerable advantage in understanding what is said, it is an advantage that is rather quickly and easily acquired" (p. 213).

Only two studies in the literature address vowel recognition specifically. Gulian & Hinds (1981) used monosyllabic words to evaluate the effects of listener experience on vowel intelligibility. Their results showed a non- significant experienced listener advantage. In this study, listeners were permitted to omit answers. An interesting and perhaps revealing finding was the large increase in the number of omissions for naive listeners. The authors interpret this finding as representative of attitude and motivational differences between the two groups (e.g. willingness to guess). The absence of a significant listener advantage may also have had to do with the extreme difficulty of the listening conditions. That is, Gulian & Hinds note that half the talkers were 'only beginning to articulate', that vowels were heard in isolated monosyllables, and that the randomized speaker design may have prevented experienced listeners from gaining an advantage by adjusting to an idiosyncratic vowel space.

Contrary to the above finding, McGarr & Gelfer (1983) showed substantial differences between listeners in percent correct vowel perception. An interesting descriptor of the differences in error patterns between the two groups of listeners was the 'spread' of errors. That is, when experienced listeners made an error, there was homogeneity in their incorrect response choice. Inexperienced listeners scattered their errors across all possible vowel categories.

The work reviewed in this section leaves many questions unresolved regarding the significance and form of the experienced listener advantage. A noteworthy trend appears to be that the perceptual advantage afforded to experienced listeners is significant only for studies using a blocked speaker design; that is all tokens of one talker are heard before any tokens of the next talker (Mangan, 1961; Markides, 1970; McGarr, 1978; McGarr & Gelfer, 1983; Thomas, 1963), but not for those using randomized speaker designs in which talkers change from trial to trial (Boothroyd et al., 1977; Gulian & Hinds, 1981). If the experienced listener advantage is dependent on repeated exposure to a single talker, then the effect may be the result of the experienced listener's ability to adjust rapidly to the idiosyncrasies of an individual talker. As discussed, this speaks directly to the notion of a stable but deviant phonology. However, there is no obvious explanation for the magnitude of the advantage changing as a function of the stimulus material (sentences vs. words vs. segmented words). One might ask whether the two groups of listeners use the same decoding strategies; that is 'look for information' in the same places. As suggested earlier, experienced listeners may learn to compensate for reduced coarticulatory information in the signal whereas inexperienced listeners do not. With respect to vowel intelligibility, it would be revealing to compare the sources of information used by listeners in determining vowel identity in the speech of deaf vs. hearing talkers. While questions about the use of coarticulatory information in normal vowel perception have been extensively pursued (see following section), no parallel research has been done regarding the perception of vowels produced by the deaf. Although it is relevant to several of the prominent theories of vowel production (and speech production in general) in the hearing impaired, we have no information as to the mediating perceptual processes used in the decoding of deaf speech.

In summary, three hypotheses regarding the underlying deviant nature and therefore the unintelligibility of vowels produced by the deaf are:

1. The Absent Target Hypothesis
2. The Deviant Phonology Hypothesis
3. The Abnormal Coarticulation Hypothesis

These hypotheses have not been adequately tested for two reasons. First, no study has assessed intrasubject variability in production. Second, no study has assessed the perceptual consequences predicted by each hypothesis.

NORMAL VOWEL PERCEPTION

Results from a number of experiments converge on the notion that the normal listener perceives speech by continuously integrating information over time. Listeners are able to integrate acoustically disparate cues into a unified percept because of their origin in a single vocal tract gesture (see Liberman., 1982 for a full review). Thus, a perceptual mechanism is proposed to capture the dynamics of the motor act. Evidence for the integration of cues over time may be found in trading relations and cue equivalence for consonant voicing (Summerfield & Haggard, 1977), place (Bailey & Summerfield, 1980), and manner (Best et al., 1981; Dorman et al., 1979; Fitch et al., 1980).

These findings are particularly interesting with regards to vowel perception since vowels are classically defined as static points in acoustic space. However, perceptual studies suggest that vocalic information is dispersed throughout the utterance. That is, the F1 and F2 steady state target frequencies may be only one of the potential sources of vocalic information in a 'normally produced' utterance.

Calibration to a Talker's Vowel Space:

An early study by Ladefoged & Broadbent (1957), using synthetic speech, manipulated the formant characteristics of a carrier phrase relative to an upcoming test vowel. Results showed predictable changes in the perception of the test vowels as a function of the acoustic characteristics of the preceding sounds. This finding supports the hypothesis that listeners focus their attention "not on the absolute values of the frequencies of the formants, but on the relationship between those frequencies and the general ranges of frequencies which seem to be characteristic of the speaker." (p.99).That is, if given the opportunity, the listener 'calibrates' to the talker's vowel space (Joos, 1948).

The perceptual advantage for vowel identification with vs. without a carrier phrase has been replicated in other studies (Dechovitz, 1977; Verbrugge et al., 1977). While Dechovitz did find greater overall error rates for vowels heard without as opposed to with a carrier phrase, he showed no significant effect of the precursor's particular acoustic characteristics. Therefore, while there appears to be vocalic information beyond the confines of the syllable, the calibration effect cannot be simply explained in terms of normalization to an F1/F2 space.

Verbrugge et al. (1976) used real speech stimuli and did not show a formant related calibration effect. In both blocked and randomized speaker designs, there were no significant differences between vowel error rates depending on whether the test vowel was or was not preceded by exemplars of the talker's point vowels. However, when stimuli were presented with erroneous cues regarding speaking rate (mismatching precursors and test syllables spoken at different rates) vowel intelligibility decreased significantly.

By manipulating the durational characteristics of a carrier phrase, Ainsworth (1974) showed that under conditions of formant ambiguity, perceived vowel duration, and therefore vowel identity, is a function of the relative duration of sounds heard just before the test vowel.

In summary, evidence is mixed with regards to the perceptual effect of a carrier phrase on vowel intelligibility. Studies which do show the effect disagree on the source of information used during the 'calibration routine'. Theories suggest calibration to the talker's vocal tract characteristics and/ or to the perceived rate of the utterance.

Consonant Context Effect:

Real and synthetic speech stimuli have been employed to study the role of transitional information in vowel perception. Experimental designs fall into three major categories:

A. Studies comparing the intelligibility of vowels produced and heard in CVC syllables to portions of the test vowels gated out of their original consonant contexts (Assmann et al., 1982; Bond, 1976,1975; Fujimura & Ochiai, 1963; Ochiai & Fujumura, 1971)

B. Studies comparing the intelligibility of vowels produced and heard in consonant context to the contextual information heard alone; that is, no steady state information (Ohde & Sharf, 1977; Strange et al., 1977)

C. Studies comparing the intelligibility of vowels produced and heard in consonant context to vowels produced and heard in isolation (Assmann et al., 1982; Diehl et al., 1981; Gottfried & Strange, 1980; Howell, 1981; Macchi, 1980; Strange & Gottfried, 1980; Strange et al., 1976,1977,1978,1979; Shankweiler et al., 1978).

Some studies combine Types A and B (Strange et al., 1978). Additionally, the effects of synthetic vs. natural speech stimuli (Diehl et al., 1981; Shankweiler et al., 1978), and speech vs. non-speech contexts (Fowler & Shankweiler, 1978; Howell, 1981; Pisoni et al., 1979) have been considered.

Studies in categories A and B are unequivocal in their findings. Information contained in formant transitions to and from a test vowel carry and in some cases are sufficient to specify vocalic identity.

Gated vowel stimuli have been either 50 ms. (Fujimura & Ochiai, 1963; Ochiai & Fujumura, 1971) or 80 ms. (Assmann et al., 1982; Bond, 1975) in duration, with rise and fall times then added to these values. Bond (1976) reiterated a single pitch pusle excised from a vowel to create an 80 ms. steady state signal. Results show from a 3% (Bond, 1976) to a 25% (Ochiai & Fujimura, 1971) increase in error rate as a function of removing contextual information. These studies confirm that 'steady state' stimuli are less intelligible than when in their original dynamic contexts and that perceptual errors tend to reflect the coarticulatory environment from which the stimuli were removed.

Type B studies present listeners with nothing but transitional information. Ohde & Sharf (1977) used real speech and presented listeners with a continuum of stimuli ranging from consonant information only (aperiodic portion of the syllable) to the full CV and/or VC syllables (aperiodic portion with transitional or transitional plus steady state information). Results show that vocalic transitions are a sufficient cue for the test vowel and that, in fact, the vowel's steady state portion may serve only as a 'redundant cue'. Strange et al. (1978) compared 'silent center' to 'center only' stimuli (with both fixed and variable durations) originally produced in a /bVb/ context. For the majority of the test speakers, the 'silent center' condition yielded the highest intelligibility scores. Comparing fixed to variable 'center only' stimuli revealed a perceptual advantage due to varying stimulus duration.

It is possible, then, for vowel identity to be completely specified by dynamic coarticulatory information. A vowel heard in consonant context is more richly specified and therefore less perceptually ambiguous than is the same vowel heard as a steady state signal (Gottfried & Strange, 1980).

The results of Type C test procedures yield a far from unanimous opinion regarding the significance of the context advantage. Here, test vowels are produced in isolation and in consonant context. This, of course, introduces differences between the stimuli beyond the presence or absence of transitional information. The equivalence of the steady state information in the two productions, and therefore the validity of the test paradigm, is entirely determined by the degree of accuracy and variability inherent in the speaker's production skills. It is for this reason that this design could not be considered to investigate the consonant context effect in the speech of the hearing impaired. One would have to wonder whether a perceptual test such as this would be evaluating anything beyond how well the talkers produced vowels in isolation vs. in context. This question, while quite interesting, is beyond the scope of this thesis. Moreover, as will be discussed below, it is doubtful that these studies actually test the perception of steady state vs. non-steady state vowels. Nevertheless, for the sake of completeness, they will be briefly reviewed here.

A number of studies have shown that errors in vowel identification were significantly higher for isolated vowels than for vowels in consonant context (Gottfried & Strange, 1980; Strange et al., 1976, 1979, 1980). This effect was stable across mixed and blocked speaker conditions, and across talkers (men, women, and children). Furthermore, it has been shown that the consonant context advantage is not determined by the phonological inappropriateness of a given token's appearance as an isolated vowel (e.g. lax vowels do not occur in isolation in English) (Gottfried & Strange, 1980; Strange et al. , 1979). Additionally, Gottfried & Strange showed that response familiarity (words vs. nonsense syllables) was of no significance. This study included an acoustic analysis (F1,F2,duration) of those tokens misidentified by a majority of listeners. It was not possible to 'predict' misidentifications from spectrographic data. The authors suggest that vowels are redundantly specified and that several parameters are important in specifying correct as well as incorrect productions.

Others have argued that the context effect is insignificant. These investigators claim that the details of the response form, the speaker-listener dialect match, and the extent of listener training have benefited vowel perception in consonant context and thus inflated the difference scores (between consonant context and vowel only conditions) in previous studies (Assmann et al., 1982; Diehl et al., 1981; Macchi, 1980).

The results of several studies suggest, however, that Type C designs do not, in fact, address the question of steady state vs. non-steady state vowel perception. A critical distinction between Type A and B studies vs. Type C studies is the use of vowels isolated as the result of waveform editing vs. naturally produced isolated vowels. Shankweiler et al. (1978) compared the perception of naturally produced isolated vowels with vowels created by iterating a single pitch pulse. Results showed twice as many errors for the latter signal. ØVE synthesized steady state vowels showed the same decrement in performance in comparison to the natural isolated vowels.

This finding suggests that naturally produced isolated vowels contain information beyond that specified by steady state formant frequency. Specifically, these stimuli may include information in natural amplitude and pitch contours, diphthongization and natural onset and offset characteristics (Assmann et al. , 1982; Strange & Gottfried, 1980; Shankweiler et al. , 1978). Assmann et al. directly tested the intelligibility of vowels produced in isolation vs. gated centers of these vowels and showed a significant increase in errors for the latter. Linear Discriminant analysis was used to obtain an index of relative 'category goodness' for individual vowel tokens. This index (reflecting a token's distance, in a multidimensional space, from the group mean for that vowel) was derived on the basis of several combinations of acoustic features measured. The correlations between this value and obtained identification scores provides a measure of the association between the acoustic parameters used in calculating the index, and those extracted by the listener. Of particular relevance here, are the results for full isolated vowels. Indices which include dynamic measures (onset and offset formant slopes) all yield significant correlations with perception. Furthermore, these correlations are higher than any obtained with indices based on steady state information alone.

The controversy over Type C studies could, then, relate largely to the character of the supposed 'isolated vowel' stimuli. Naturally produced isolated vowels seem to represent some kind of a perceptual mid-ground. Under conditions of some uncertainty (e.g. dialect, rate, formant ambiguity) adding consonant contexts will improve intelligibility. However, when naturally produced isolated vowels already include some dynamic information, it is possible, if conditions are well-controlled, for these stimuli to be as intelligible as their context-laden counterparts.

It seems quite reasonable to assume that Type A and B studies do, in fact, compare the intelligibility of steady state vs. non-steady state vowels. Comparisons between truly static tokens and those 'normally' produced as part of an articulatory gesture, uniformly reveal a perceptual advantage. Moreover, in light of the finding that vowels can be perceived in the absence of steady state information, it is difficult to suppose that dynamic coarticulatory information is not used, to some extent, in specifying vowel identity in running speech. This is not to suggest that steady state vowels are unintelligible. Instead, it appears that if given the opportunity, the perceptual system will benefit from information regarding dynamic vocal tract gestures.

OVERVIEW

Those studies which have asked questions about the underlying nature of vowel production (and speech production in general) in the hearing impaired, have offered several theories to account for their findings. Three major theoretical positions may be summarized as follows:

1. Hearing-impaired speakers do not have precise representations of articulatory/acoustic vowel targets. That is, they are somewhat random in their employment of articulatory gestures for any given vowel (Absent Target Hypothesis).

2. Hearing-impaired speakers, as a result of limited auditory input, have abstracted different phonological systems than the norm. With regards to vowels specifically, they operate within a deviant phonological space. The parameters upon which vowels are differentiated may be unique but are well-controlled nonetheless (Deviant Phonology Hypothesis).

3. A major problem in the speech of the hearing-impaired can be described as a 'breakdown in motor programming'. The rules which govern coarticulatory behavior in normal speech production are disrupted. The most significant errors in their speech rest in the timing and/or movement domains. This aspect of their speech may prevent the listener from decoding their utterances in the normal way (Abnormal Coarticulation Hypothesis).

One major question to be answered, then, is whether hearing-impaired talkers show acoustic evidence of well-controlled articulatory targets in vowel production. Although there are studies of intrasubject variability in first and second formant frequencies for vowels produced by normal hearing adults (Broad, 1976; Pisoni, 1976; Peterson & Barney, 1952; Potter & Steinberg, 1950) and children (Eguchi & Hirsh, 1967; Kent, 1978; Lieberman, 1980) there has been no systematic exploration of this variable in the speech of the hearing impaired. Similarly, there has been no work done to substantiate the claim that there are features of vowel production, other than the F1/F2 space, which differentiate between vowels and are produced with a high degree of precision. That is, no study has fully tested the notion of a deviant phonology regarding vowel production by the hearing impaired.

While numerous objective measurements support the notion of reduced or abnormal coarticulatory effects in the speech of the hearing impaired, no work has been done to substantiate this claim with perceptual measures. In 'normal vowel perception', at least 3 sources of information have been shown to be potential carriers of vowel quality:

1. steady state formant frequencies
2. transitional information in the surrounding articulatory gesture
3. information available in a carrier phrase (re. rate and/or vowel space)

If, in fact, hearing-impaired speakers provide less coarticulatory information to the listener and/or operate within a well-defined idiosyncratic vowel space, these aspects of production should reveal themselves in a controlled perceptual task of the sort just reviewed. Furthermore, this type of an experimental design would provide a forum for replicating Calvert's (1961) findings as well as for investigating the particulars of the 'experienced listener advantage' with regards to vowel perception.

FOOTNOTES

[1] In this text, 'incorrect' refers to those tokens which are not perceived as the phoneme intended by the talker.

[2] The 'range' in this study is defined as the difference, in Hz, between the highest and lowest mean value that an acoustic parameter obtains across all the test vowels.

[3] Range is defined here as it was in Angelocci et al. (1964).

CHAPTER III

METHODS

TALKERS

Eight high school students served as talkers for this study. Six of the children were hearing impaired and two had normal hearing. The six hearing-impaired talkers all had a congenital severe-to-profound hearing loss and no other major handicapping condition. Each was considered to be of at least average intelligence as determined by school records. All came from homes in which the primary spoken language was English. At the time of the study, all six were enrolled in an oral school for the deaf in the New York metropolitan area. The two hearing talkers had normal speech production skills and served as age-matched controls.

Further restrictions were placed on the choice of speakers since a number of 'talker related' factors are known to decrease the accuracy of the acoustic measurements which formed the data for this study.

For instance, we know that as the fundamental frequency increases, the number of harmonics which define each formant decreases, as does the accuracy of formant measurements (Lindblom, 1962; Monsen & Engebretson, 1983; Peterson & Barney, 1952). Therefore, due to a characteristically high f0, it can be difficult to estimate formant frequency in the speech of young children. Additionally, improper control of laryngeal and/or velopharyngeal structures may lead to a variety of spectral abnormalities which can make formant measurements problematic (Stevens, 1977; Stevens et al., 1983). Such source function problems are common in the speech of the deaf. Difficulties in making accurate estimates of formant frequency for this population have been noted (Monsen, 1976a; Stein, 1980; Stevens et al., 1983; Suonpaa & Aaltonen, 1981).

Therefore, to enhance the accuracy of the acoustic measurements made in this study, the talkers were (1) limited to males who were considered to be post-pubescent and consequently had low fundamental frequencies [4] and (2) free of any major phonatory problems (e.g. inability to maintain phonation, extreme breathiness, falsetto voice, frequent pitch breaks) as judged by their speech teachers.

An attempt was made to include deaf children who had speech production skills representative of a wide range of overall intelligibility [5]. Informal judgements regarding intelligibility (below average, average, above average) were made by each of the student's speech teachers. In order to substantiate these judgements a recording was made of each student reading a list of 20 sentences ranging in length from 5 to 11 words. These sentences were developed by Smith (1972) and were composed of words which were within the vocabulary of the deaf children. Following the procedures used by Smith, these recordings were audited by inexperienced listeners each of whom heard only one child. The listeners were instructed to write down as much as possible of what was said and a percent intelligibility score for key words was obtained. Table 3.1 shows the scores of the six hearing-impaired children who served as talkers for this study, along with their ages and pure tone averages in the better ear. The ages of the two hearing talkers are also shown.

Based on the results of this sentence test, Smith (1972) divided her hearing- impaired subjects into quartiles of intelligibility (first quartile = most intelligible, fourth quartile = least intelligible). Talkers D3 and D4 would have been ranked in the top of the first quartile, D2 in the bottom of the first quartile, D1 and D5 in the top and bottom of the second quartile, respectively, and D6 in the third quartile.

LISTENERS

50 experienced and 50 inexperienced listeners participated in this study. Listeners were considered experienced if they had been working with the hearing impaired on a day-to-day basis for a period of 2 to 10 years. All experienced listeners were employed at schools for the deaf and the vast majority were classroom and/or speech teachers. None of the experienced listeners was employed at the school attended by the talkers in this study.

Listeners were considered to be inexperienced if they reportedly had 'no contact with hearing-impaired individuals on other than an incidental basis'.

Table 3.1

Age, PTA and Intelligibility of Talkers

TALKER	AGE (yrs)	PTA(dBHL)	Smith sentences (% correct)
D1	16	107	23
D2	16	88	32
D3	18	95	63
D4	15	85	72
D5	18	93	13
D6	17	105	9
H1	15	--	--
H2	14	--	--

All inexperienced listeners were undergraduate college students. All listeners were native speakers of American English.

STIMULUS MATERIALS

To determine if decisions about vowel quality in the speech of hearing-impaired and/or normal hearing children are made with reference to the talker's vowel space, all of the test vowels were spoken in the context of a carrier phrase. Following the format used in the study of this phemonemon in 'normal vowel perception' (Ainsworth, 1972; Dechovitz, 1977; Ladefoged & Broadbent, 1957; Vebrugge et al., 1976) the carrier phrase was designed to be short, simple and to include exemplars of the talker's point vowels: /i, ɑ, u/. The carrier phrase was:

"You got me the .....".

To evaluate whether listeners use dynamic vocal tract information vs. steady state information alone when identifying vowels produced by hearing-impaired and/or normal hearing children, the test vowels were spoken in a CVC context. A /bVb/ context was chosen for two reasons:

1. This consonant context allowed the experimenter to verify (visually) that the talker occluded his vocal tract in the production of the CVC

sequence. This was critical to the validity of the perceptual experiment. That is, if closure were omitted, it would be impossible to address questions regarding a listener's ability to extract transitional information from the signal.

2. This is the phonetic context most frequently used in studies which examine the importance of transitional information in vowel perception.

The following seven vowels were chosen for this study:

/i/ /I/ /ɛ/ /ɑ/ /ʌ/ /ʊ/ /u/

It should be noted that the test word appears at the end of the utterance. It was hoped that this design would make the waveform editing process less difficult. In hearing talkers, increased vowel duration has been noted for phrase final syllables, a durational effect called 'prepausal lengthening' (Klatt, 1975, 1976). While there is some evidence to suggest that hearing-impaired talkers also show this effect (Reilly, 1979), it was possible that some of the children in this study would not. Klatt (1976) has noted, however, that "durations are lengthened relative to the inherent duration for the segments in question" (p.1220). That is, prepausal lengthening does not disturb the relative duration of vowels within a given talker's system. Therefore, we may reasonably assume that the presence or absence of the effect would not influence measures of within-subject duration as a function of vowel type.

In order to measure intrasubject variability in vowel production a large number of repetitions of each vowel was needed. Studies of this aspect of vowel production in hearing talkers have used from 5 (Eguchi & Hirsh, 1969) to 20 (Pisoni, 1976) repetitions. Fifteen tokens of each vowel by each speaker were obtained yielding a total of 105 tokens per talker (7 vowels x 15 repetitions) and 840 tokens in all (105 tokens x 8 talkers). The order in which the vowels were spoken by the children was randomized within and across speakers [6].

INSTRUMENTATION AND RECORDING SESSIONS

Recordings were made using a TEAC A-7030U stereo tape recorder and a Realistic electret lapel microphone. The frequency response of the microphone coupled to the recorder was measured using a Bruel & Kjaer sweep frequency sine random generator (#1024), microphone amplifier (#2603), and graphic level recorder (#2305). A block diagram of the calibration system and the obtained frequency response are shown in Figures 3.1 and 3.2, respectively.

A Bolt, Beranek & Newman Model 501 accelerometer, P18 power supply and conditioning module were used to record nasal vibrations (Stevens et al., 1975). The accelerometer was attached to the lateral surface of the nose and its output fed into the second channel of the TEAC recorder. The accelerometer weighs 1.8 grams and did not appear to interfere with the children's speech production in any way. It was hoped that physical

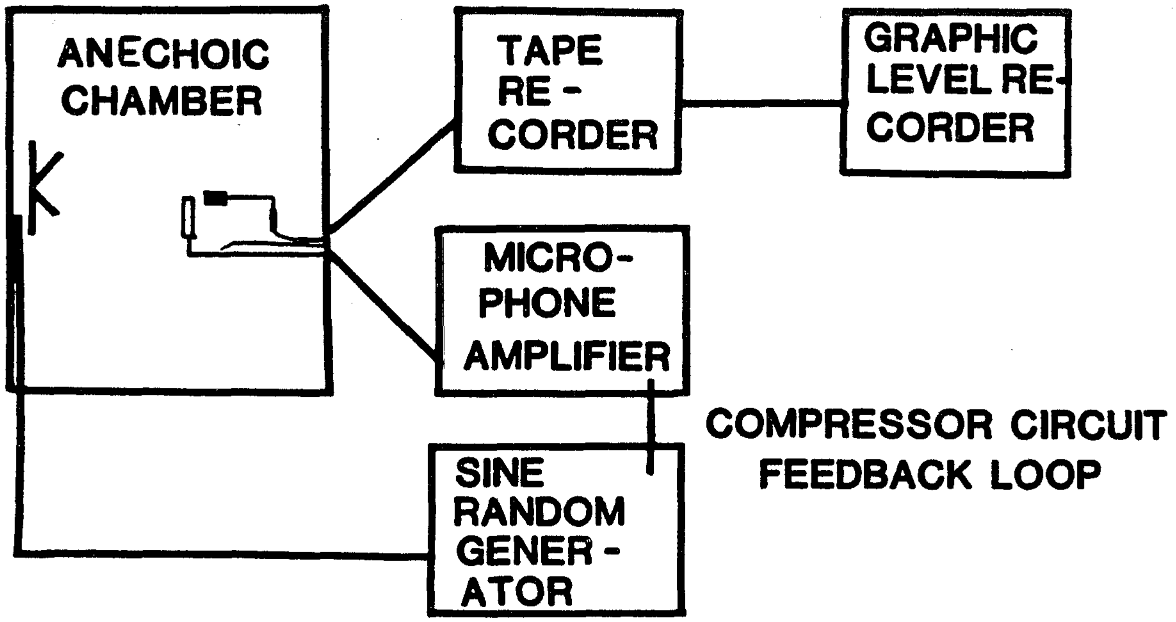

Figure 3.1 Block Diagram of Calibration System.

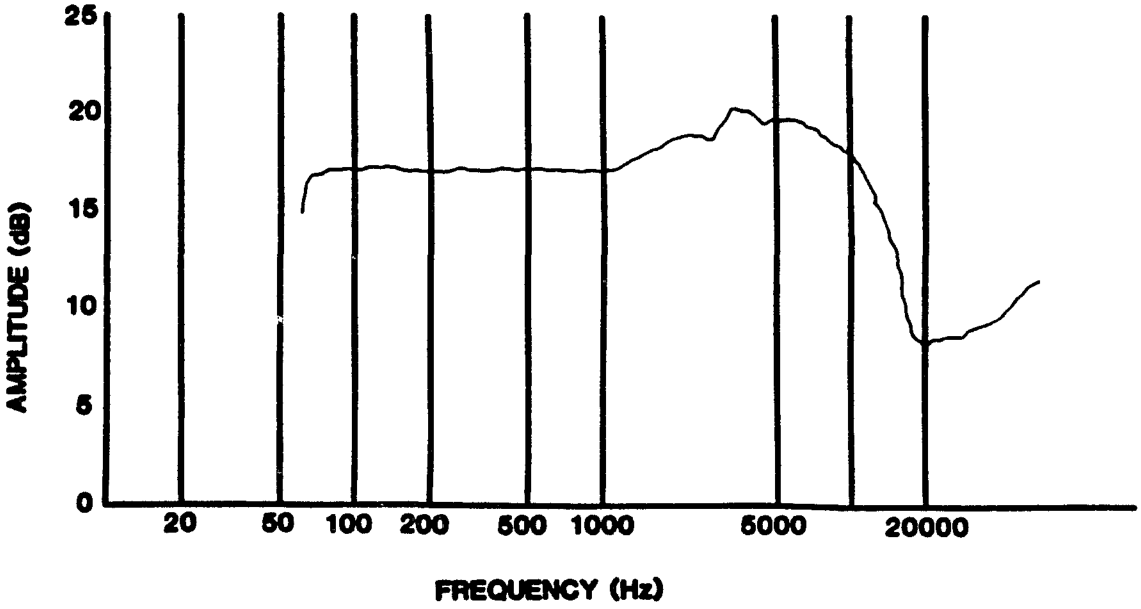

Figure 3.2 Frequency Response of Recording Equipment.

measurements indicating the presence or absence of nasality would be helpful in attempts to discriminate between oral and nasal formants in the acoustic analysis (to be discussed below). Moreover, these measurements might have been informative had we encountered difficulty in predicting the intelligibility of tokens on the basis of formant frequency.

Recordings were made in a quiet room at the school that the hearing-impaired children attended. Each recording session had the following format:

The child was familiarized with the test vowels in the /bVb/ context and with their orthographic representation. Symbols recommended by the children's speech supervisor were used and are shown below with their IPA counterparts:

ē	i	ė	ä	u	u̇	ü
/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/

Alternative stimulus cards had been prepared in case a child did not appear to be familiar with a particular symbol. When this happened, the child was given the opportunity to look at and read the alternatives. The symbol with which he appeared to be most familiar was used during the test proper. For example, a number of the talkers seemed to prefer the symbol "oo" for /u/, "ee" for /i/, "eh" for /ɛ/ and so on.

The child was asked to say the /bVb/ nonsense syllables several times in random order. This practice session continued until it was clear that the child and the experimenter agreed on the vowel sound to be associated with each symbol. When necessary, the children were given a real word in which a vowel appeared. This was to assure that measurements of vowel production skill did not, in fact, reflect a child's problem with the orthography. The children also practiced reading the /bVb/ syllables in the context of the carrier phrase. During the test proper, the child read each of the seven test syllables preceded by the carrier phrase a total of 15 times. The order in which the vowels were spoken was randomized as previously described.

A child was asked to repeat an utterance if (1) he showed disapproval of his own production, (2) having previously demonstrated the ability to utter the test vowel unambiguously, it was clear to the experimenter that the child had produced a different vowel from his repertoire or (3) the initial or final consonant in the test syllable was omitted (as judged visually by the experimenter not observing bilabial closure). Recording sessions lasted approximately 35 to 45 minutes.

SEGMENTATION PROCEDURES

Using a synchronous dichotic input program (12 bit A/D converter) available on the Haskins Laboratories VAX 11/780, the two channel recordings were digitized at a sampling rate of 10 kHz per channel, using a Nyquist filter (upper cutoff frequency: 4900 Hz, 40 dB/octave slope) and high frequency pre-emphasis. To eliminate low frequency room noise, frequencies below 100 Hz were filtered out using two Allison Variable Filters (Model 2AB) in series. Acting as a high pass filter, with a cutoff frequency of 97.5 Hz and a 60dB/octave slope, this system yielded a maximum signal-to-noise ratio of 46dB.

Once digitized, the sentences produced by the children were segmented in accordance with 3 sequential rule governed procedures. The purpose of this multi-staged segmentation protocol was to generate 3 types of stimuli from a single 'parent' utterance. These were:

1. Sentence stimuli
2. Word stimuli
3. Gated Vowel stimuli

(all 3 versions having identical steady state formants but differing with regards to the amount and type of surrounding contextual information).

Segmentation was accomplished by means of a Haskins Laboratories waveform editing and display program ('WENDY'). This editor provides continuously variable cursors (head and tail markers) which may act simultaneously on one channel or more. Thus, it was possible to extract both the speech file as well as the corresponding 'nasality file'. Segmentation was done by hand by the experimenter. Success at each stage of segmentation was determined both visually (by observing the waveform) and auditorily (by listening to the edited sample).

The first segmentation protocol extracted a single sentence (carrier phrase and test syllable) from the speech signal. In this case, the rules were simply to set the head and tail markers as close as possible to the onset of 'You' and offset of the /bVb/ syllable without eliminating any articulatory/acoustic information. 840 sentence files were created (8 talkers x 7 vowels x 15 repetitions).

The second segmentation protocol extracted the /bVb/ test syllable from each sentence file. Tail markers were not altered in any way. Head markers were adjusted to a point just before the onset of the bilabial burst. Without exception, this was a very simple task for the tokens produced by the hearing-impaired children. As has been noted before (Osberger & Levitt, 1979) hearing-impaired speakers often insert within-phrase and/or within-sentence pauses. This finding was confirmed here since the resolution of individual words within each sentence was quite clear. The task was somewhat more laborious for tokens produced by the hearing children since the separation between words was less easily determined.

The final segmentation procedure isolated a 100 ms. steady state portion of the test vowel. The two primary objectives in this gating procedure were the exclusion of CV and VC transitional information and the creation of a 100 ms. stimulus. A segment was considered to be free of transitional information if it excluded any increases or decreases in amplitude associated with oral opening or closure, respectively, and if the period of the waveform in the isolated segment remained constant [7]. Once a steady state portion of the vowel had been delineated, head and tail markers were set to +50 ms. around a center point and the segment was extracted.

For some vowels, there was evidence of erratic change in amplitude throughout the center portion of the waveform. This instability was clearly unrelated to movement towards or away from a consonantal closure. In these cases, the most steady state portion of the vowel was located and the extracted file simply included ±50 ms. around this point.

Finally, a linear ramping function was imposed on these gated vowel stimuli to produce 10 ms. rise and fall times. This was done to eliminate transients at the onset and offset of the segmented tokens.

Sample waveforms of each of the 3 stimulus types spoken by a normal hearing and deaf talker are shown in Figures 3.3 and 3.4, respectively.

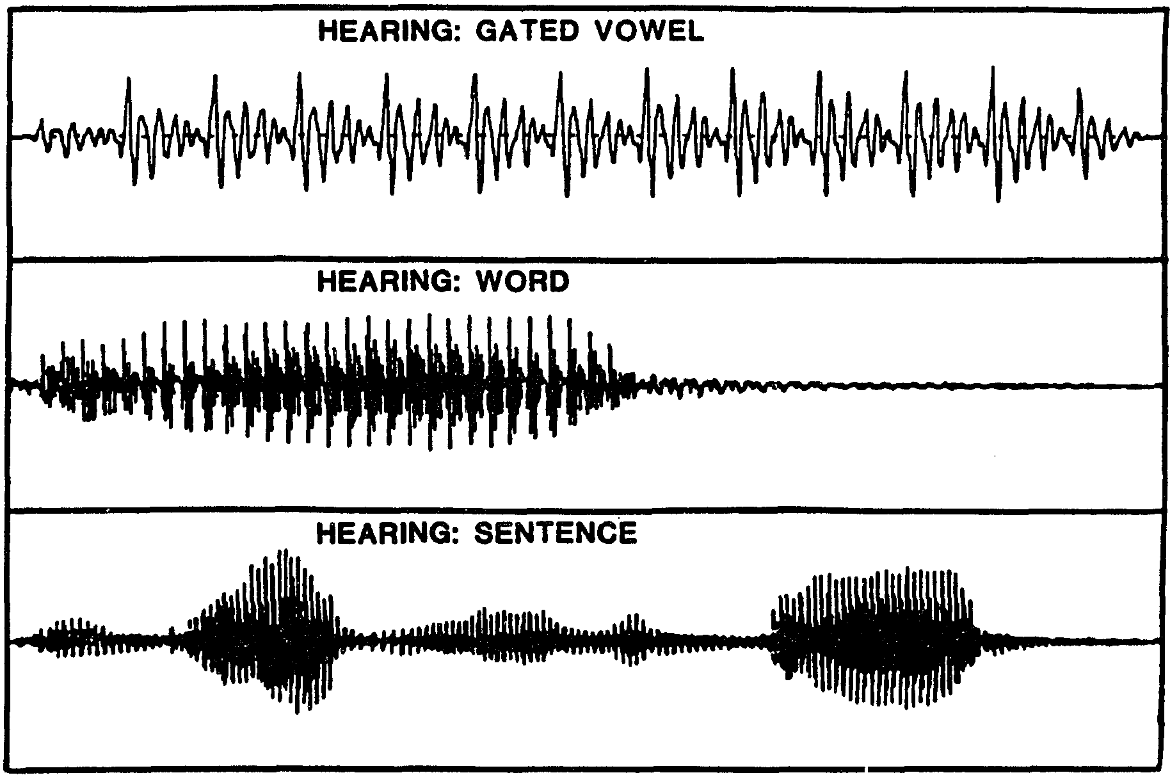

Figure 3.3 Sample Waveforms of Sentence, Word and Gated Vowel Stimuli: Hearing Talker.

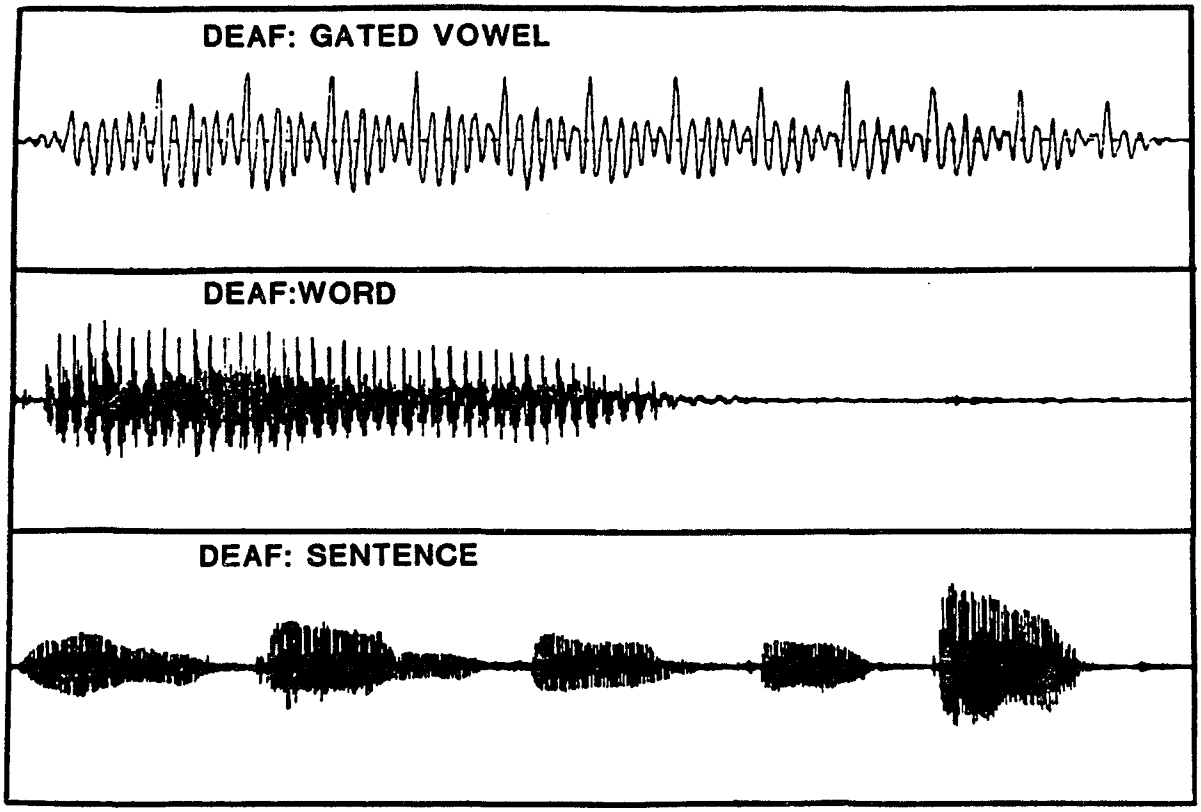

Figure 3.4 Sample Waveforms of Sentence, Word and Gated Vowel Stimuli: Hearing-Impaired Talker

DESIGN OF PERCEPTUAL EXPERIMENT

The sentences, words, and gated vowels created by the segmentation procedure served as stimuli for a perceptual experiment. Listeners were asked to identify the vowel produced and the hearing status of the talker for each token under the three conditions noted above.

In the Gated Vowel condition, then, listeners were provided with steady state formant and f0 information alone. In the Word condition, the same vowel was heard with the addition of information available in the surrounding articulatory gesture. In the Sentence condition, listeners were given the opportunity to sample the talker's vowel space and/or other talker related information available in a carrier phrase (e.g. rate). Since each parent utterance appears in every condition, changes in vowel intelligibility across conditions can be ascribed to the use of one of the aforementioned sources of non-steady state information in vowel perception. Similarly, it should be possible to delineate the acoustic basis upon which 'deaf voices' are identified as such by looking at differences in ability to identify the hearing status of a talker as a function of condition.

2520 stimuli (840 tokens x 3 stimulus types) were converted to analog form. Five listening tapes were developed. There were 168 stimuli per condition per tape (3 tokens x 7 vowels x 8 talkers). Each token was separated by an interstimulus interval of 5 seconds. The stimuli were blocked into groups of 24. Each group was separated by an intergroup interval of 7 seconds. Listeners heard all of one condition before they heard any of another. Talkers were randomized within conditions. Each condition was preceded by a practice set consisting of 14 items of the appropriate stimulus type (each vowel spoken once by a normal hearing and hearing-impaired child not used in the study). The practice sets were identical on all five tapes. 7 seconds separated the practice set from the test proper. Leader tape separated listening conditions. The order of presentation of conditions (Sentence, Word, Gated Vowel) was randomized across the five tapes.

10 experienced and 10 inexperienced listeners heard each of the 5 tapes. Testing was done in a quiet room using a Pioneer RT1050 tape recorder and Sennheiser headphones (Models 424X and 414X). The frequency response of the recorder did not vary from manufacturer's specifications by more than .8dB at any frequency.

As mentioned earlier (see discussion of Type C studies, Chapter II) the details of the listener's response form in this kind of a perceptual study were not a trivial matter. The vowel symbols chosen were those thought to elicit the most unambiguous interpretation [8]. They are shown below with their IPA counterparts:

ee	i	eh	ah	u	uh	oo
/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/

To reduce possible contamination of listeners' responses by a mismatch between stimulus and response form, answer sheets differed across stimulus conditions. That is, for every condition there was a complete match between the phonetic context in which the vowels were heard and their written symbolic representation. 'Task differences', then, should have been effectively eliminated. Samples of each of the response forms are shown in Appendices A1 - A3.

For each trial, listeners were required to: (1) mark the symbol that best described the vowel sound they heard and (2) to indicate whether they thought that the speaker was hearing or hearing impaired. They were asked not to leave any answers blank and encouraged to guess if necessary. A full set of instructions is shown in Appendix B. The entire listening session lasted one hour.

During the instructional period, great care was taken to see that the listeners were familiar with the vowel sounds and their orthographic symbols. The experimenter read all the symbols aloud in the /bVb/ context as well as the isolated vowel context. In addition, listeners read and heard every vowel in a /kVd/ context. This context provided listeners with a sample of each vowel in a 'real' word and was, at the same time, an equal phonetic mismatch for all test stimuli. A practice session followed during which listeners had to demonstrate their ability to associate correctly the vowel sounds with their orthographic symbols.

Some listeners had particular difficulty remembering the symbols for /U/ vs. /ʌ/. In these cases, they were allowed to write the words 'could' and 'cud' above the appropriate columns on their response forms.

ACOUSTIC ANALYSIS

Vowel duration measurements were made from the segmented /bVb/ files. The onset of the vowel was defined as the onset of periodic activity in the waveform following the release of the initial consonant. The vowel's offset was defined as the offset of periodic activity in the waveform or the point at which the periodic activity associated with the vowel reached its minimum amplitude, whichever came first. In this way, voiced transitions into and out of the vowel were considered part of the vowel.

Formant frequency, bandwidth, and amplitude were estimated using a Linear Predictive Coding analysis program at Haskins Laboratories. The LPC program used the autocorrelation method to calculate reflection coefficients on a frame by frame basis. Fundamental frequency estimates were based upon a modified cepstral processing technique.

LPC analysis has been widely used in measuring normal speech. This procedure has been shown to be at least as accurate as standard spectrographic techniques in extracting the first two vowel formants. In general, the measurement error for F1 is larger than that for F2 (Monsen

& Engebretson, 1983). These investigators have also shown that LPC analysis is extremely accurate (within .1 Hz) in the measurement of fundamental frequency. This technique has also been used successfully in analyzing the speech of clinical populations (Bush, 1981; Gerratt, 1983; Levitt & Stromberg, 1972).

While the sampling rate was set to 10 kHz, other user-controlled parameters were manipulated to arrive at a 'best fit' for the source characteristics of the given data set. The length of the analysis window was set to 20 ms. with a 10 ms. step size between adjacent frames. The number of predictor coefficients was set to 16. Using fewer than 16 coefficients often led to estimates which appeared to 'miss' one of the first two formants. More than 16 coefficients did not appear to increase the accuracy of the analysis.

LPC analysis was performed on the 100 ms. gated vowel files (before the rise and fall times had been adjusted). This produced 9 frames of analysis for each token. For the majority of cases, the two lowest formants selected by the program were of reasonably steady frequency from onset to offset and were of appropriate bandwidth and relative amplitude. For these tokens the values for the two lowest formants in the centermost frame were taken as F1 and F2. A decision was made not to average over adjacent frames in order to keep the time domain of the analysis at least somewhat similar to that represented by a narrowband spectrographic section.

In a number of cases, the program misidentified F1 as F2 and/or F2 as F3. Occasionally spurious peaks in the spectrum and/or nasal resonances were mistakenly identified as oral formants. In these instances, if an appropriate and stable F1 and F2 could be traced, by hand, throughout the LPC printout, formant values near the center of the utterance were extracted. Phonetic transcription by an experienced listener, the extracted nasality files, and a set of criteria regarding relative amplitude and bandwidth of allowable formants were used in determining what represented a spurious peak, nasal resonance, or misidentified formant. In most instances, the nasality files were uninformative. That is, the presence or absence of nasality was seemingly unrelated to the program's ability to identify appropriate formants.

There remained 177 cases (21%) for which further analysis was required in order to make unambiguous estimates of F1 and F2. These tokens were analyzed on a Kay Elemetrics Digital Sonograph (7800). Wideband spectrograms and spectrographic sections were produced. An amplitude rule was used to determine the general location (within 50 Hz) of the first two formants. These estimates served as a kind of template in interpreting the results of the corresponding LPC analysis. That is, for these special cases, values of F1 and F2 frequency were taken to be those in the LPC analysis which most closely matched the estimates derived from spectrographic sectioning. Fundamental frequency estimates were always reasonable and no additional analysis techniques were required.

FOOTNOTES

[4] An additional reason for choosing talkers in this age group is that they would be beyond the age range during which intrasubject variability in vowel production could be ascribed to developmental processes (Eguchi & Hirsh, 1969).

[5] The perceptual part of this study compares the intelligibility of vowels presented with various amounts of dynamic vocal tract information. Little insight would been gained, then, from talkers whose vowels were completely unintelligible even in the most ideal conditions. Therefore, when the speech teachers chose potential candidates for this study, they were asked not to choose students whose speech was considered to be completely unintelligible.

[6] 5 randomized lists of the test vowels were generated (A-E). 3 randomized groupings of lists A-E were produced (1-3). The order in which groups 1-3 were read by the children approximated a Latin square design.

[7] In a number of cases, the test vowels were too brief to be able to meet these criteria. That is, isolating a 100 ms. portion of the vowel necessitated the inclusion of what appeared to be CV and/or VC transitional information. These tokens were not included in the analyses of the perceptual effects of listening condition.

[8] An informal survey of naive subjects suggested that these symbols were most likely to be 'read' correctly.

CHAPTER IV

RESULTS

ACOUSTIC ANALYSIS OF VOWELS PRODUCED BY HEARING AND HEARING-IMPAIRED TALKERS

Overview:

Every vowel token produced was submitted to an acoustic analysis in which the following four physical parameters were measured:

1. first formant frequency
2. second formant frequency
3. fundamental frequency
4. vowel duration

The purpose of this analysis was to address two major theoretical issues. The first was that of intrasubject variability. It was hoped that measuring this aspect of production, in conjunction with the standard 'range of mean values', would be more informative in describing and categorizing a talker's vowel system than either measure alone. In addition, direct comparisons were made between normal hearing and hearing-impaired talkers regarding token-to-token variability in each acoustic domain. This provides relative information regarding articulatory control and allows us to consider the possible substitution of control among parameters.

The second issue at hand, not unrelated to the first, was that of a 'deviant but well-defined phonological space'. The question asked is whether there exists a previously unexplored combination of articulatory/acoustic parameters which serves to distinguish between vowel targets in these talkers in a way that the standard F1 and F2 frequencies cannot (irrespective of accuracy within the normal vowel space).

Hearing vs. Hearing-Impaired Talkers as Groups:

Mean values of F1, F2, f0, and duration as a function of vowel target were determined for the hearing-impaired and normal hearing talkers as groups. Group means were calculated by averaging the mean values for individual talkers. These values are shown in Table 4. 1. Also shown are the ranges of each measure across vowel targets. For each variable, range was determined by calculating the difference between the highest and lowest mean value obtained across the seven test vowels. A comparison between the two groups shows that the hearing-impaired talkers have a larger range of f0 values across vowel targets than the normal hearing talkers (19 Hz vs. 12 Hz, respectively). The same is true for vowel related changes in

Table 4.1

Hearing and Hearing-Impaired Group Means and Ranges as a Function of Vowel Target: F1, F2, f0 and Duration

	HEARING IMPAIRED (n=6)				NORMAL HEARING (n=2)			
	f0	F1	F2	DUR	f0	F1	F2	DUR
/i/	154	324	2142	232	141	276	2407	205
/I/	150	459	1944	196	141	456	1971	143
/ɛ/	151	434	1987	205	140	570	1783	162
/ɑ/	135	786	1287	303	134	799	1249	241
/ʌ/	144	633	1346	179	140	622	1127	164
/ʊ/	150	500	1308	184	146	449	1037	146
/u/	154	402	1393	252	146	342	890	209
RANGE	19	462	855	124	12	523	1517	98

duration (124 ms. vs. 98 ms., respectively). The opposite trend was noted for F1 and F2 range. The normal hearing talkers showed ranges of 523 Hz and 1517 Hz for F1 and F2, respectively. Comparative values for the hearing- impaired group as a whole were 462 Hz and 855 Hz, respectively. In the hearing-impaired group, the lowest mean value of second formant frequency was for the vowel /ɑ/. The F2 range would have been reduced by between 21 and 106 Hz if /ʊ/ or /u/ had been arbitrarily chosen to estimate the lowest second formant (as if often done whn analyzing the vowel space of a normal hearing talker).

Mean values of duration and f0 across vowel targets were determined for the two groups as a whole and are shown in Table 4.2. On the average, vowels produced by the hearing impaired were 40 ms. longer than vowels produced by the normal controls. Mean fundamental frequency was 7.1 Hz higher in the hearing impaired as compared to the normal hearing group. Also shown in Table 4.2 ar mean values for individual talkers. All but one hearing-impaired subject (D6) reflect the group mean in that their overall vowel duration was longer than the mean values of either normal hearing talker. This is not true for f0. Mean fundamental frequency for 3 out of 6 hearing-impaired talkers (D2, D4, D5) falls within the range covered by the two normal controls.

Table 4.2

Hearing and Hearing-Impaired Mean Values of Duration and f0

Across Vowels: Group and Individual

	Dur (ms)	f0 (Hz)
D1	211.3	169.1
D2	269.4	146.0
D3	208.2	151.0
D4	226.8	138.2
D5	238.1	133.7
D6	174.7	152.0
H1	183.5	135.6
H2	179.2	146.8
Mean HI	221.4	148.3
Mean H	181.4	141.2
HI range	175-269	134-169
H range	179-184	136-147

Individual Data:

Mean values of F1, F2, f0, and duration measures as a function of vowel target are shown for individual talkers in Tables 4.3-4.6, respectively. Also shown are the talker's individual ranges for each of these measures. Individual ranges vary greatly from the group means reported in the previous section. That is, for f0, F1, and duration, the hearing-impaired talkers vary among themselves to the extent that they demonstrate ranges both 'less than or approximately equal to' and 'greater than' those of the normal controls. Within-group F2 ranges for the hearing impaired, while always less than those of the controls, vary by as much as 288 Hz. Also noteworthy is that with the exception of D4, the hearing-impaired talkers all demonstrate ranges which are from 283 to 418 Hz less than the smallest range of the normal talkers.

In general, the ranges produced when averaging across hearing-impaired talkers do not adequately represent the skills of individual talkers. Intersubject differences are too large to warrant general statements about these deaf speakers as a group.

To determine the distinctive properties of individual vowel systems, each talker was analyzed separately. Vowel differentiation based on F1 and F2 values was considered first.

Figures 4.1-4.8 show the mean F1/F2 values of each vowel target for each talker and will be called 'mean plots'. A talker's vowel production skills would typically be determined on the basis of this information. Figures 4.9-4.16 show the F1/F2 plots for each token produced by each

Table 4.3

Mean Values of F1 (Hz) as a Function of Vowel Target and Ranges:

Individual Talkers

	/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/	Range
D1	280.4	338.9	360.0	846.7	758.4	345.5	321.8	500
D2	427.3	499.8	505.1	763.2	603.9	514.7	501.6	336
D3	356.7	390.8	423.2	788.0	546.5	518.5	427.3	498
D4	260.5	458.0	475.0	760.7	616.2	604.3	237.8	566
D5	268.1	470.0	322.4	766.3	627.5	451.6	435.6	431
D6	349.5	598.9	518.1	789.3	646.3	567.5	451.8	440
H1	245.1	462.9	546.7	731.0	618.9	480.3	325.8	561
H2	306.1	449.8	593.8	866.7	624.3	417.4	358.5	486

Table 4.4

Mean Values of F2 (Hz) as a Function of Vowel Target and Ranges:

Individual Talkers

	/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/	Range
D1	2132.7	2137.5	2125.3	1426.9	1288.1	990.5	1011.2	880
D2	2094.3	1899.7	1967.3	1160.3	1294.2	1195.0	1225.7	934
D3	1848.1	1854.5	1849.5	1082.6	1105.9	1052.1	995.7	909
D4	2108.3	1879.6	1812.3	1228.1	1455.3	1544.1	2054.1	1147
D5	2498.4	2415.1	2386.3	1588.6	1751.9	1841.3	1779.5	859
D6	2171.7	1473.8	1781.5	1235.2	1178.1	1222.3	1293.4	994
H1	2546.7	2022.3	1771.4	1173.5	1129.1	1000.2	788.7	1758
H2	2267.7	1920.4	1794.3	1323.8	1125.0	1073.9	991.0	1277

Table 4.5

Mean Values of f0 (Hz) as a Function of Vowel Targets and Ranges:

Individual Talkers

	/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/	Range
D1	175.5	175.6	178.5	143.5	162.9	179.2	168.7	35.7
D2	148.1	147.8	145.9	143.9	145.0	146.0	146.4	4.2
D3	153.0	154.8	154.8	138.3	149.7	153.0	154.0	16.5
D4	145.8	144.5	141.4	123.7	135.5	132.1	144.3	22.1
D5	138.8	132.9	132.5	117.9	125.7	140.8	147.7	29.8
D6	163.6	147.3	151.9	141.5	146.4	148.7	164.6	23.1
H1	134.3	135.1	135.0	128.7	136.1	140.5	139.5	11.8
H2	148.1	147.3	144.4	138.9	143.1	151.8	153.2	14.3

Table 4.6

Mean Values of Duration (ms) as a Function of Vowel Target and Ranges: Individual Talkers

	/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/	Range
D1	183.2	215.4	208.1	276.1	139.4	151.2	306.0	166.6
D2	274.0	252.6	254.0	355.6	235.5	231.8	282.3	123.8
D3	246.7	159.2	178.3	305.6	148.7	152.2	266.6	156.9
D4	255.2	199.8	182.0	275.9	211.3	199.8	263.5	93.9
D5	244.4	199.1	239.2	328.2	199.8	230.5	225.7	129.1
D6	187.5	148.5	166.2	276.8	136.9	138.1	168.9	139.9
H1	214.9	139.8	156.6	243.3	164.2	146.7	218.8	103.5
H2	195.5	146.5	168.1	238.1	162.8	145.3	199.1	92.8

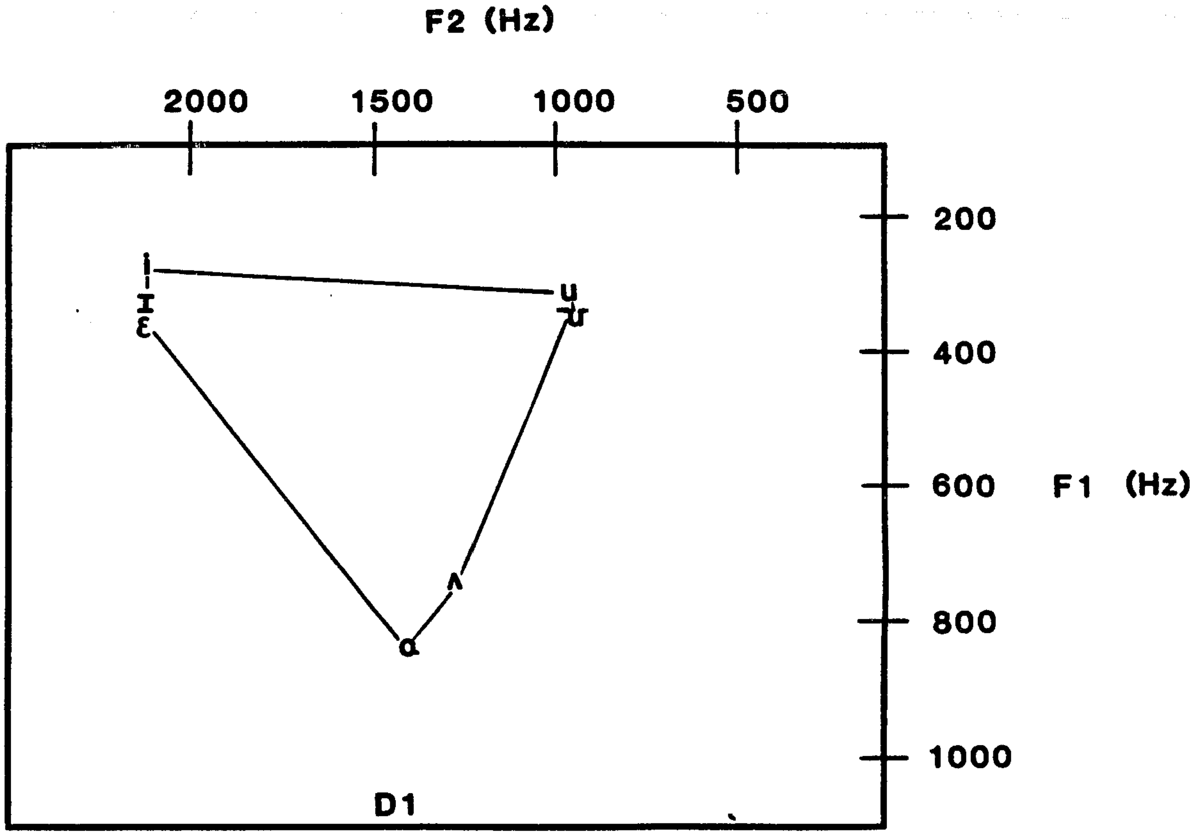

Figure 4.1 Mean F1 vs. F2 Plot: D1.

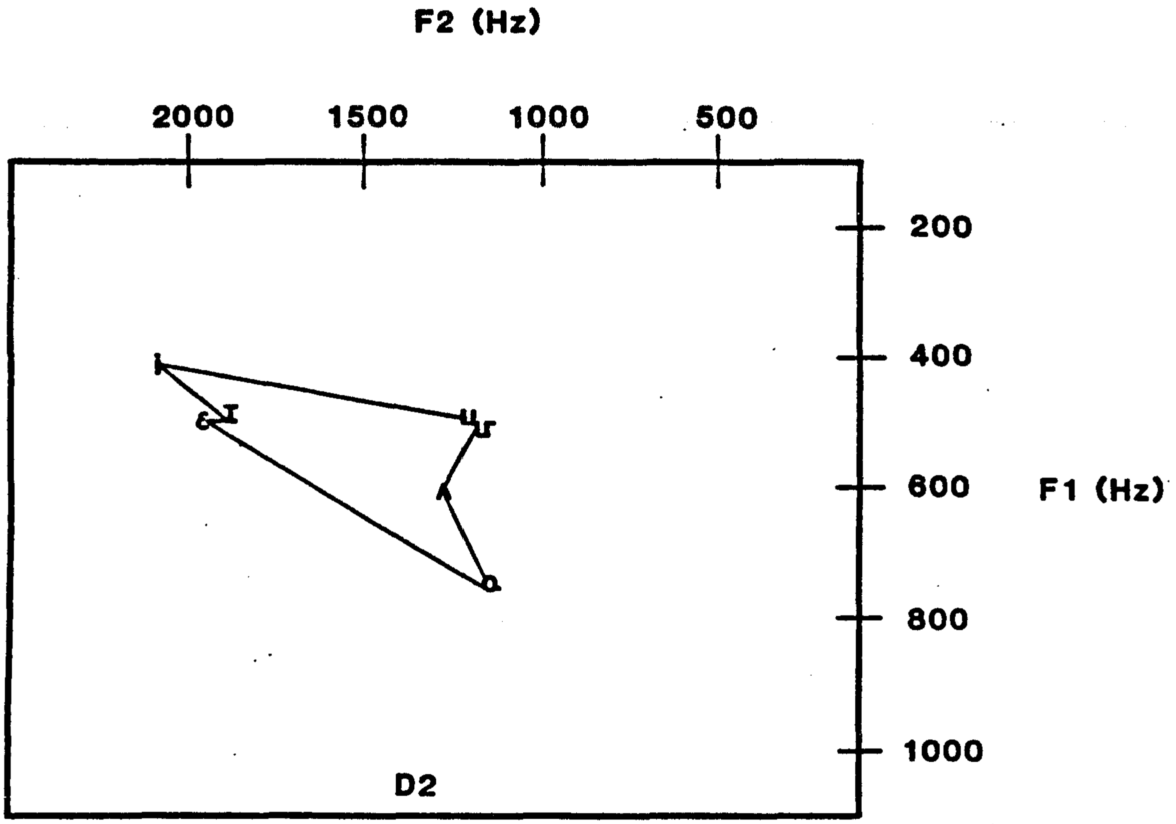

Figure 4.2 Mean F1 vs. F2 Plot: D2.

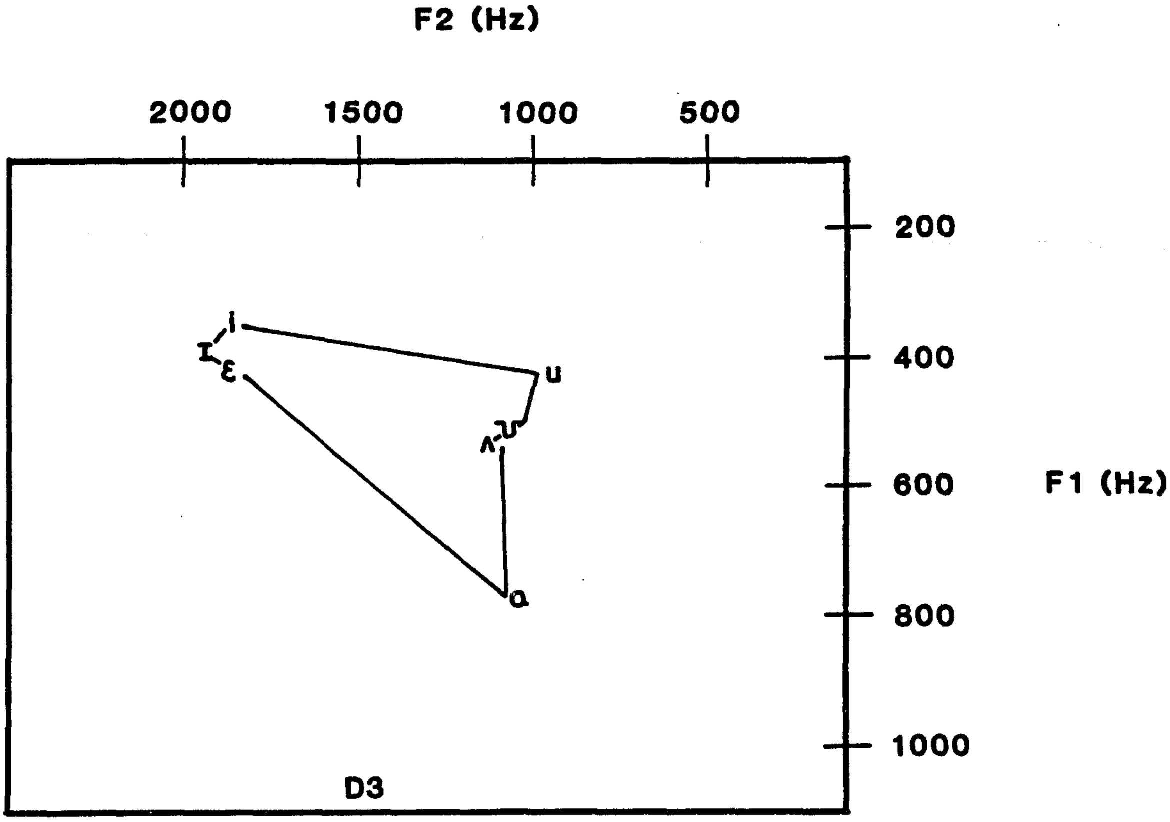

Figure 4.3 Mean F1 vs. F2 Plot: D3.

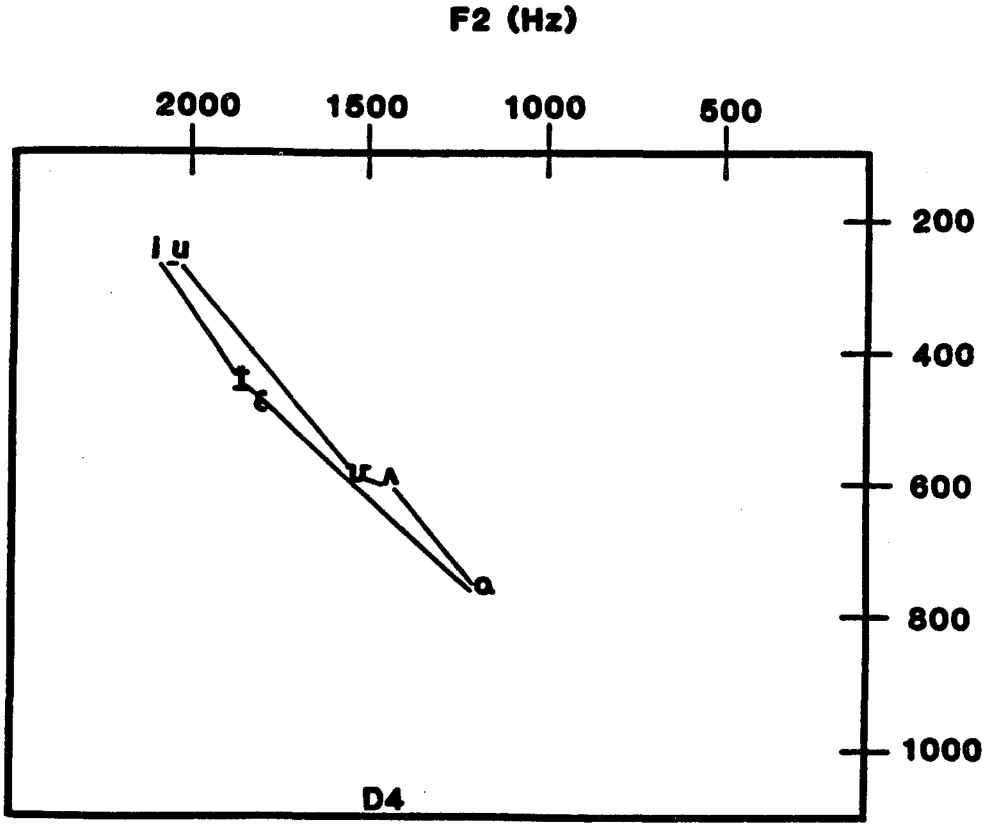

Figure 4.4 Mean F1 vs. F2 Plot: D4.

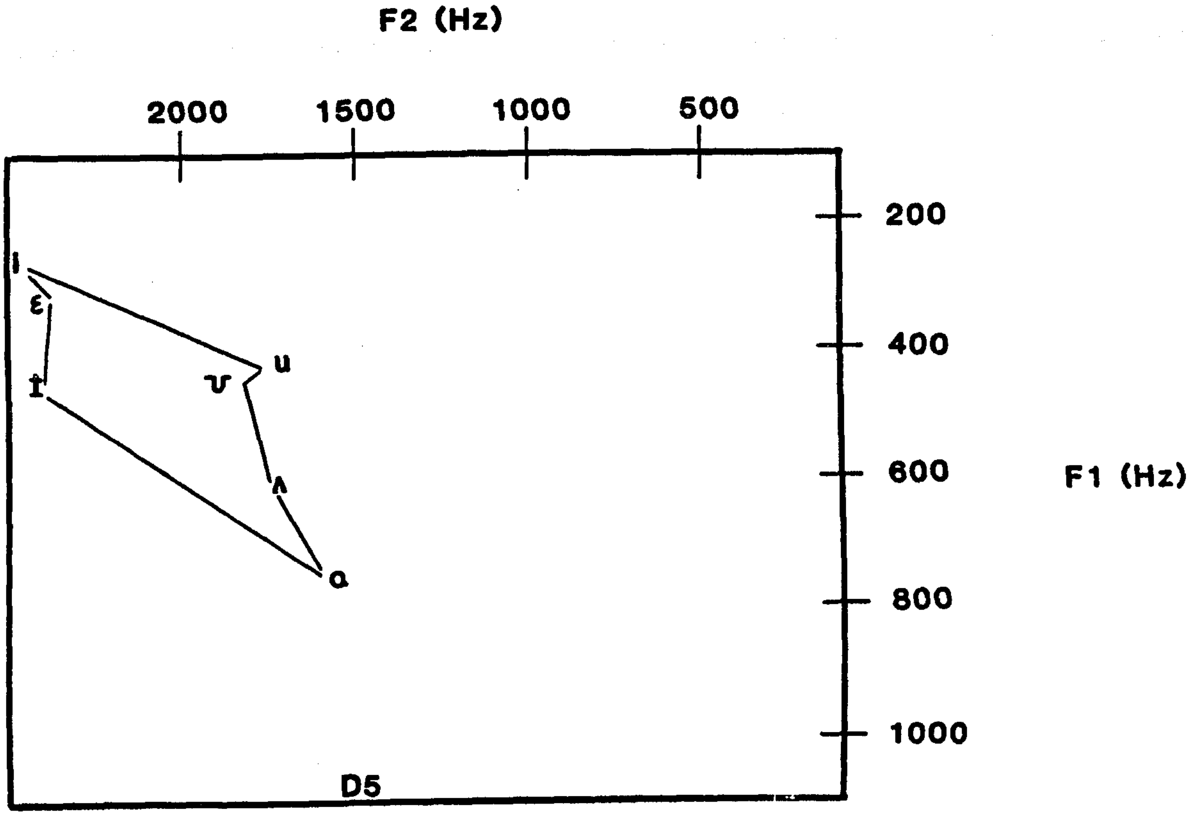

Figure 4.5 Mean F1 vs. F2 Plot: D5.

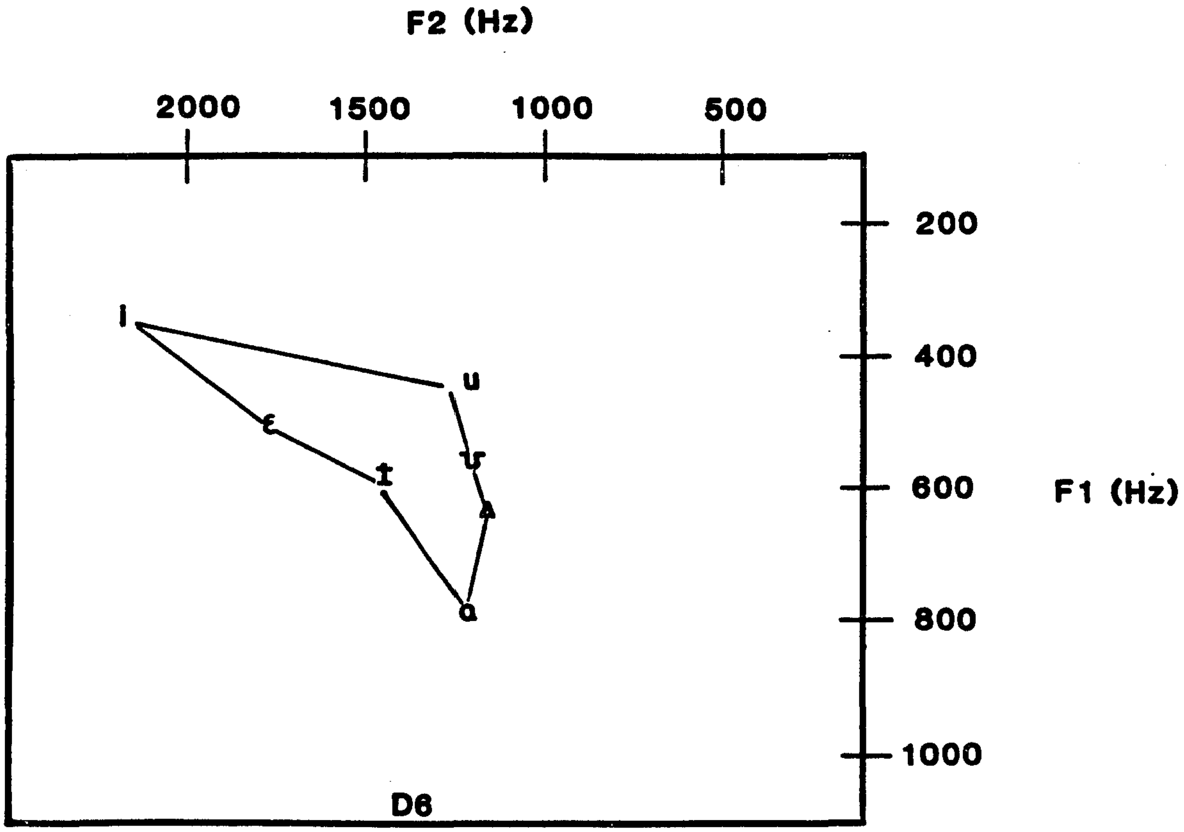

Figure 4.6 Mean F1 vs. F2 Plot: D6.

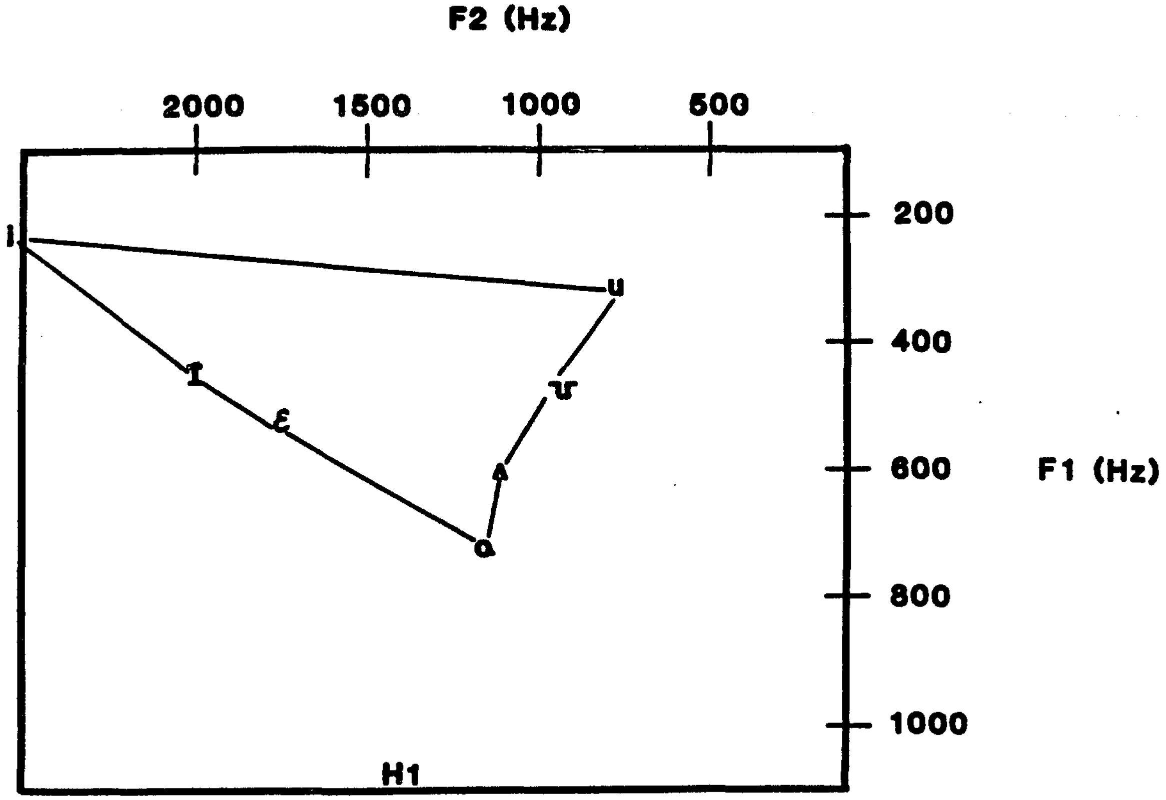

Figure 4.7 Mean F1 vs. F2 Plot: H1.

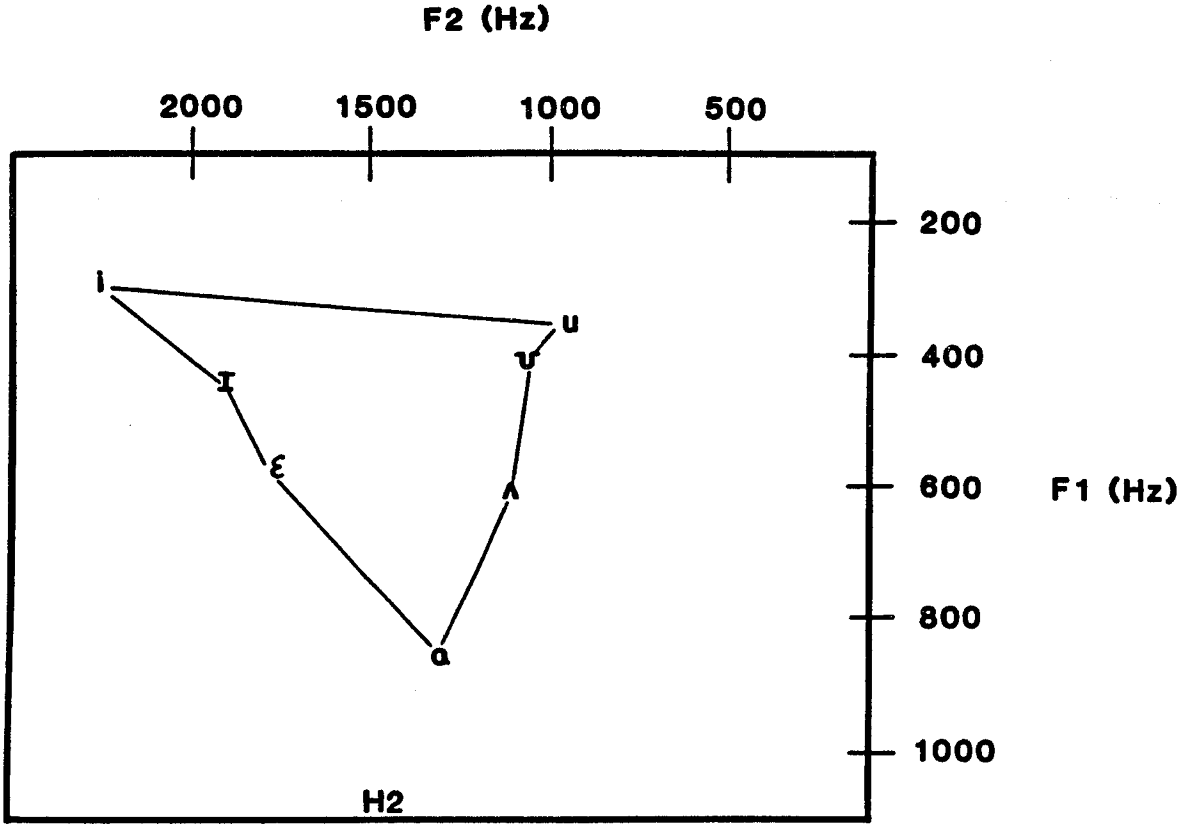

Figure 4.8 Mean F1 vs. F2 Plot: H2.

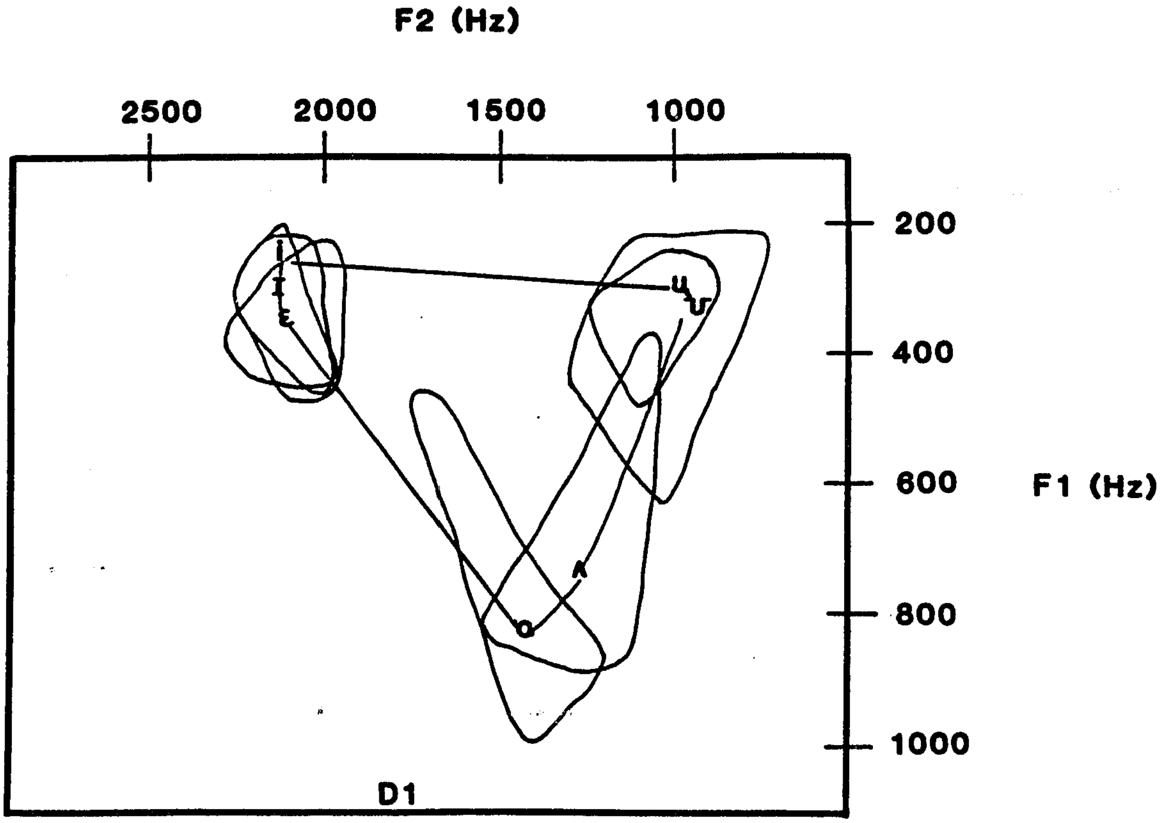

Figure 4.9 Token F1 vs. F2 Plot: D1.

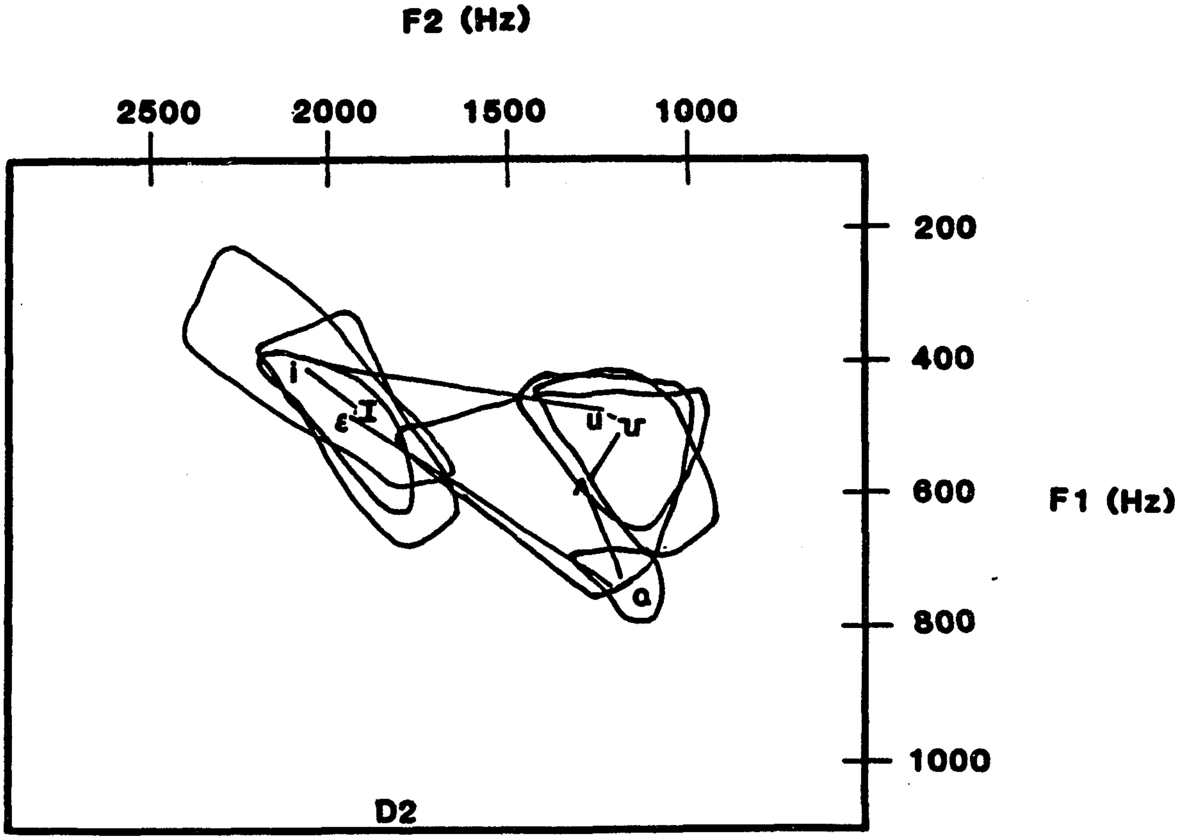

Figure 4.10 Token F1 vs. F2 Plot: D2.

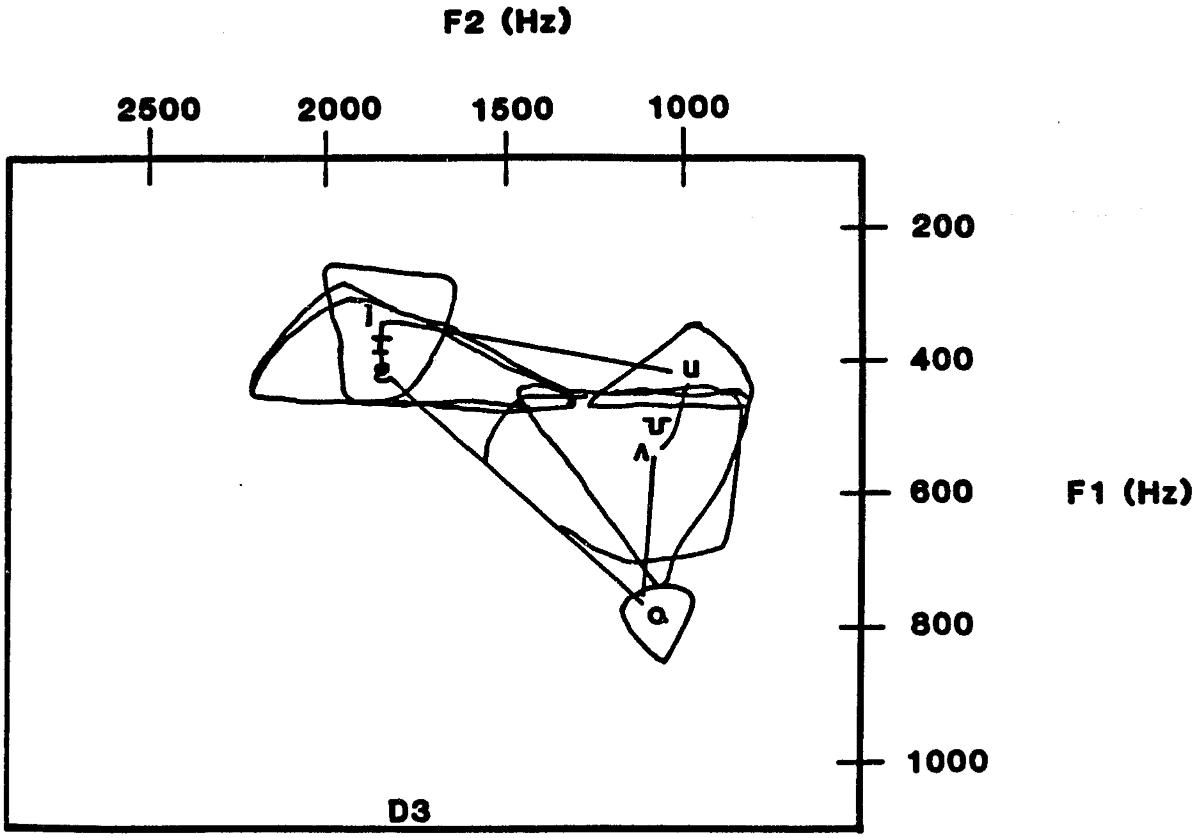

Figure 4.11 Token F1 vs. F2 Plot: D3.

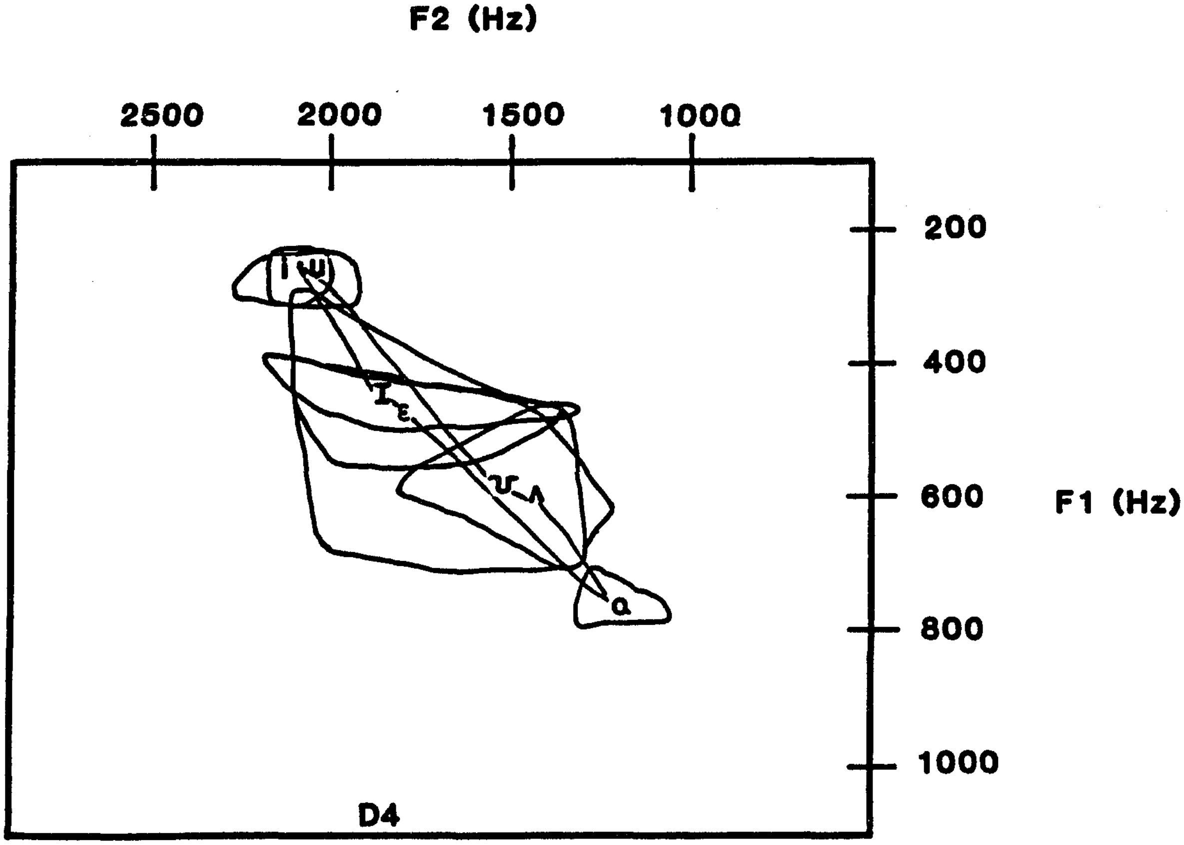

Figure 4.12 Token F1 vs. F2 Plot: D4.

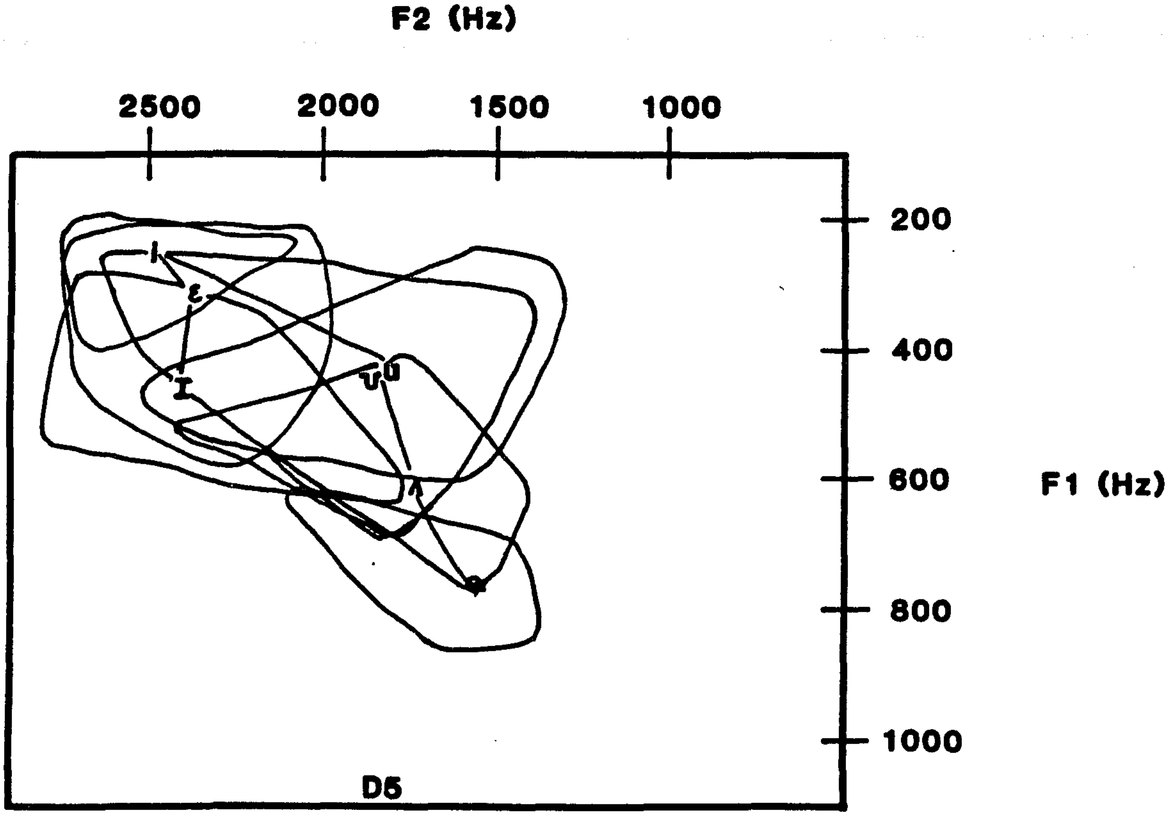

Figure 4.13 Token F1/F2 Plot: D5.

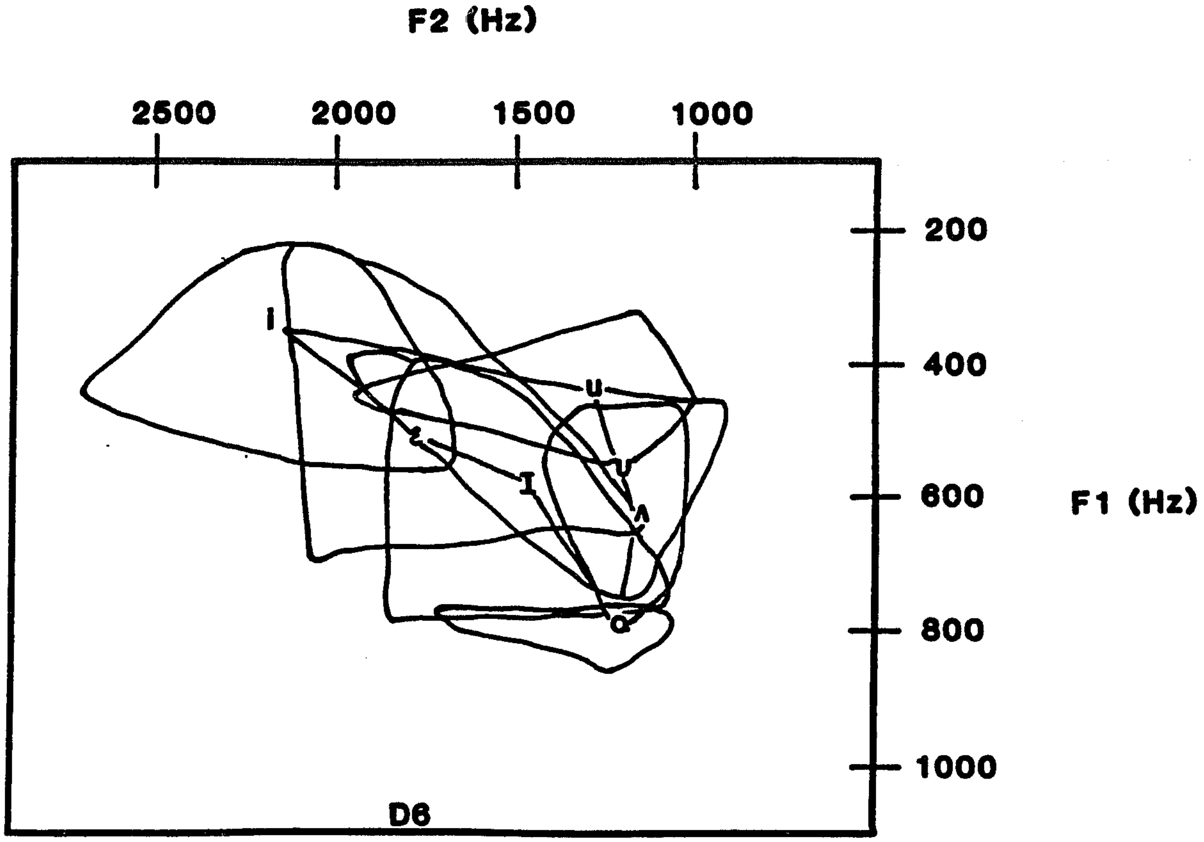

Figure 4.14 Token F1/F2 Plot: D6.

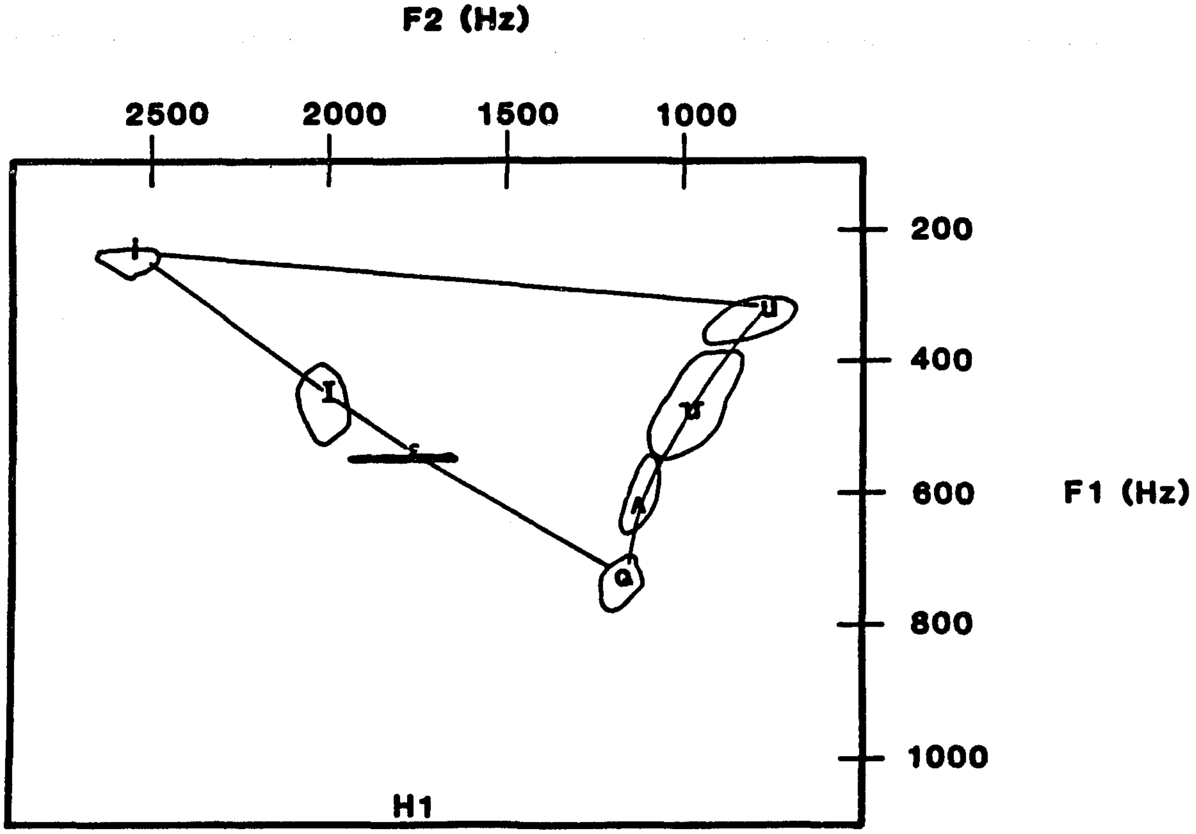

Figure 4.15 Token F1/F2 Plot: H1.

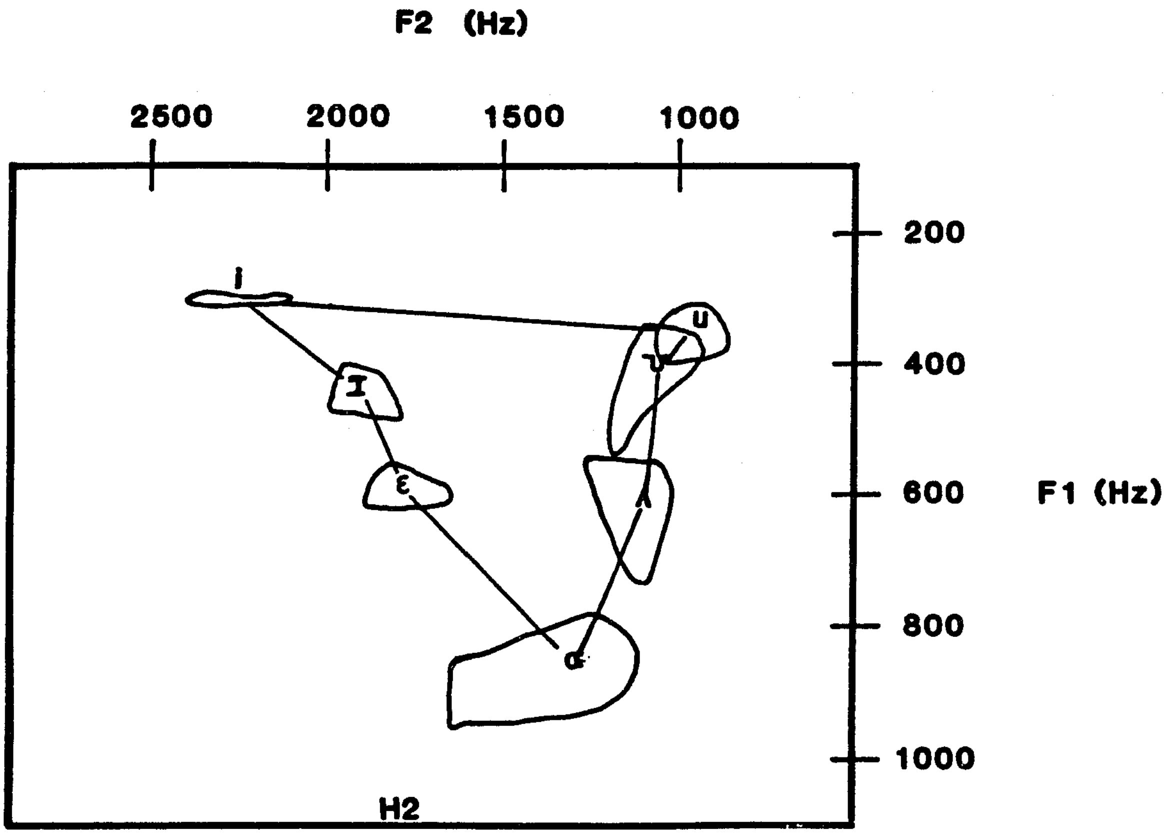

Figure 4.16 Token F1/F2 Plot: H2.

talker and will be referred to as 'token plots'. Lines have been drawn to enclose the points which represent repeated attempts to produce the same target vowel. A comparison of the two figures for each talker is revealing.

Talker D6, for example, has the second largest intervowel F2 range of the hearing-impaired group. This information alone might suggest that this talker uses well-controlled tongue movemements in the front/back dimension to differentiate between vowel targets. However, the 'token plot' for this talker (Fig. 4.6) reveals an excessive amount of token-to-token variability for each vowel and a high degree of overlap between vowel categories. Thus, while mean F2 values would suggest a large and therefore distinctive vowel space, a closer look reveals unclear vowel targets based on this parameter.

Another difference between token and mean plots is that the latter provides no comparative information regarding the articulatory control for individual vowels. The results of talker D4 highlight this deficit. Looking first at the mean data (Fig. 4.4), one would assume that the targets /i/ and /u/ were undifferentiated on the basis of F1 and F2 values as were those for /ʊ/ and /ʌ/. More importantly, the implication would be that the articulatory control used in the production of these two vowel pairs was similar as well. Inspection of the token plot (Fig. 4.12) shows an important difference between the production of these vowel pairs. That is, the vowel spaces for /ʊ/ and /ʌ/ are large, suggesting poor articulatory control in the production of these vowels. Quite the opposite is true for /i/ and /u/. Here, the F1/F2 spaces are small suggesting well-controlled movements in their production. The disparities between these two pairs of vowel targets in token-to-token variability and therefore articulatory control could not have been determined from mean values alone.

Therefore, the information to be gained from measures of intrasubject variability is far more revealing in terms of 'vowel targets' and 'vowel systems' than is the information regarding acoustic means and ranges alone.

Types of Vowel Systems:

The token plots suggest that the six hearing-impaired talkers may be categorized as follows:

A. Those who demonstrate 'Point Vowel Systems'
B. Those who demonstrate 'Front/Back Vowel Systems'
C. Those who demonstrate 'Overlapped Vowel Systems'

The following two aspects of production serve to characterize each vowel system:

1. Intravowel variability across vowels
2. The acoustic dimension of greatest formant overlap

A. POINT VOWEL SYSTEM:

The talkers in this group (D3, D4) demonstrate clearly defined and well-controlled targets for /i/, /ɑ/, and /u/. Variability increases dramatically for the non-point vowels. The difference in control for point vs. intermediate vowel targets characterizes this system. Overlap in F2 frequencies is larger than that for F1. That is, intermediate vowels are differentiated to some extent on the basis of F1. Their production is poorly controlled and nondistinctive in terms of F2. Therefore, the acoustic dimension of greatest formant overlap is second formant frequency.

B. FRONT/BACK VOWEL SYSTEM:

The talkers in this group (D1, D2) are distinctive in two ways. First, the seven test vowels are approximately equivalent with regards to variability in production. That is, in contrast to the Point Vowel talkers, subjects D1 and D2 appear to produce all vowels with approximately equal articulatory precision. Second, overlap between vowel targets is almost exclusively a result of first formant variability. That is, unlike talkers D3 and D4, the front/back dimension is well-controlled and clearly preserved whereas fine distinctions in vowel height are not. It should be noted that variability in the front/back dimension for talker D2 is not noticeably smaller than that seen in the Point Vowel talkers. In spite of this, subject D2 rarely if ever crosses the 'acoustic midline' separating front from back vowels. The consequence of this is an increased F2 range with the front/back distinction preserved.

C. OVERLAPPED SYSTEM:

The talkers in this system (D5, D6) are like the Front/Back talkers in that none of the test vowels enjoys any special status in terms of articulatory control. All vowel targets are produced with approximately equal imprecision (although the point vowels are slightly less variable than the lax vowels in F1 frequency). They are unlike any of the other talkers in that they appear to show overlap between vowel targets in both F1 and F2 frequency. Neither of the two formants adequately delineates vowel targets.

Thus, information about variability in F1 and F2 frequency for each vowel target proves to be a particularly useful tool for describing the styles of vowel production seen in these talkers. Moreover, this aspect of production appears to be an appropriate grouping variable. Figure 4.17 shows the mean standard deviations of F1 vs. F2 for each talker. Plotted this way, the 8 speakers divide themselves into the 3 categories just described (A, B, C) plus the normal controls (H). Of the two formants, variability in F1 seems to be the more discriminating variable regarding group identity. Categorizing subjects according to production style also describes the talkers in terms of overall intelligibility.

According to the background data presented in Chapter III, the talkers in Group A are the two most intelligible, those in B are in the middle, and those in Group C are the two least intelligible talkers. It should be noted that this categorization scheme is based solely on variability in production and does not consider the accuracy of vowel targets with reference to the normal vowel space.

Intrasubject Variability in F1, F2, f0, and Duration: Hearing vs. Hearing- Impaired Talkers

To compare directly articulatory control in the hearing vs. hearing-impaired groups, and to consider the importance of f0 and duration in vowel production, the standard deviations for each of the four acoustic measures as a function of vowel target were determined and are shown for individual talkers in Tables 4.7-4.10. Mean standard deviations for the four subject groups just described are plotted in Figure 4.18. These values for tense as compared to lax vowels are shown in Figure 4.19 [9]. Figures 4.18 and 4.19 confirm the impressions given by the token plots regarding patterns of F1 and F2 variability across the hearing-impaired groups. They also compare graphically, formant variability in hearing-impaired vs. hearing talkers. As can readily be seen, all three hearing-impaired groups are more variable in F1 and F2 frequency than the hearing controls. Variability in f0 increases monotonically across the four subject groups whereas variability in duration does not.

Table 4.7

Standard Deviations as a Function of Vowel Target for

Individual Talkers: F1 (Hz)

	/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/
D1	59.8	72.1	63.6	122.7	124.0	116.0	67.6
D2	89.8	106.8	93.3	29.4	102.9	73.4	81.2
D3	73.7	66.7	63.2	22.9	100.6	85.3	42.5
D4	26.9	28.1	48.8	25.3	81.0	109.7	26.4
D5	58.7	97.2	112.8	60.2	99.1	134.6	104.0
D6	134.4	124.5	135.0	19.3	106.9	118.9	50.6
H1	9.9	28.4	8.0	30.0	30.0	51.2	19.1
H2	4.8	22.5	24.6	34.7	51.0	51.5	31.0

Table 4.8

Standard Deviations as a Function of Vowel Target for Individual Talkers: F2 (Hz)

	/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/
D1	82.1	91.1	134.9	106.5	140.9	138.6	94.2
D2	172.0	136.2	149.5	62.6	229.1	157.8	155.5
D3	116.8	209.6	201.9	56.9	200.9	193.8	120.3
D4	68.8	252.9	271.1	76.0	170.9	277.0	73.0
D5	234.3	254.3	268.6	176.2	270.3	360.5	399.4
D6	297.5	273.6	289.5	171.9	107.3	224.1	321.2
H1	44.7	37.7	90.0	31.0	36.7	77.1	56.7
H2	81.0	51.8	56.4	176.6	59.8	67.9	58.7

Table 4.9

Standard Deviations as a Function of Vowel Target for Individual Talkers: f0 (Hz)

	/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/
D1	9.0	11.9	9.5	10.1	11.2	11.4	9.2
D2	7.3	4.3	4.3	4.1	3.7	5.5	6.3
D3	5.0	3.9	4.4	5.4	7.5	4.8	4.4
D4	7.7	10.5	10.0	7.4	6.2	8.4	7.1
D5	10.5	9.1	9.5	7.3	12.1	5.6	11.7
D6	8.2	9.5	6.8	7.7	8.0	10.2	10.3
H1	4.9	5.3	3.9	4.5	5.2	4.2	3.2
H2	5.2	3.8	3.1	2.8	4.2	5.7	4.3

Table 4.10

Standard Deviations as a Function of Vowel Target for Individual Talkers: Duration (ms)

	/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/
D1	38.0	58.6	57.9	35.3	22.0	34.9	62.4
D2	37.4	51.0	36.1	25.6	35.1	30.1	45.0
D3	39.8	25.3	43.9	25.9	36.2	25.7	86.7
D4	36.2	35.5	25.5	28.1	38.8	40.5	41.3
D5	62.5	27.7	45.9	43.0	37.3	31.7	49.3
D6	17.8	26.9	17.6	25.4	20.0	19.6	34.8
H1	18.6	11.8	20.6	14.5	13.5	10.7	22.8
H2	14.2	13.3	23.7	17.3	14.5	27.2	33.6

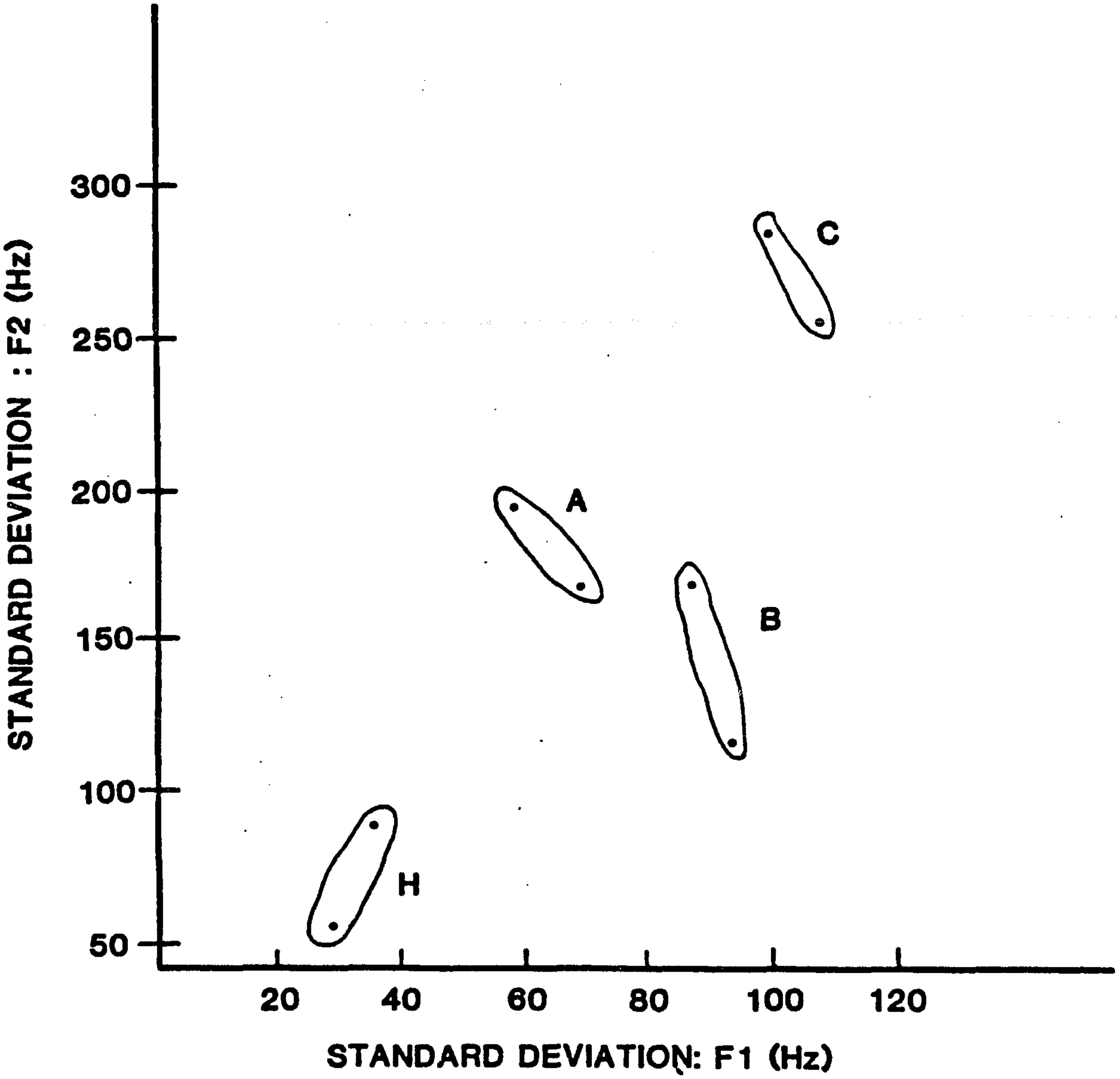

Figure 4.17 F1 vs. F2 Standard Deviations for Individual Talkers.

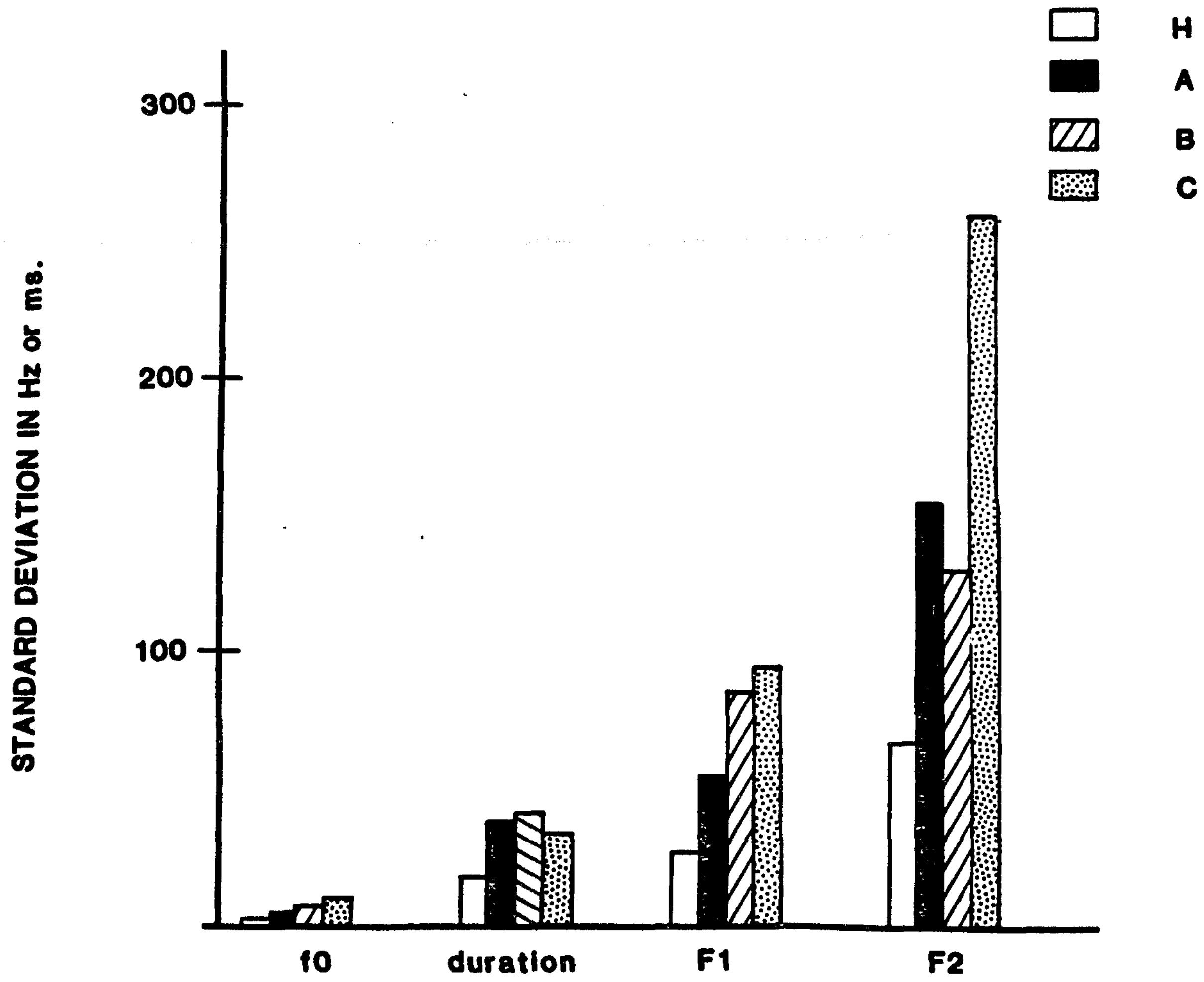

Figure 4.18 Mean Standard Deviations for Groups H, A, B and C: f0, Duration, F1, F2.

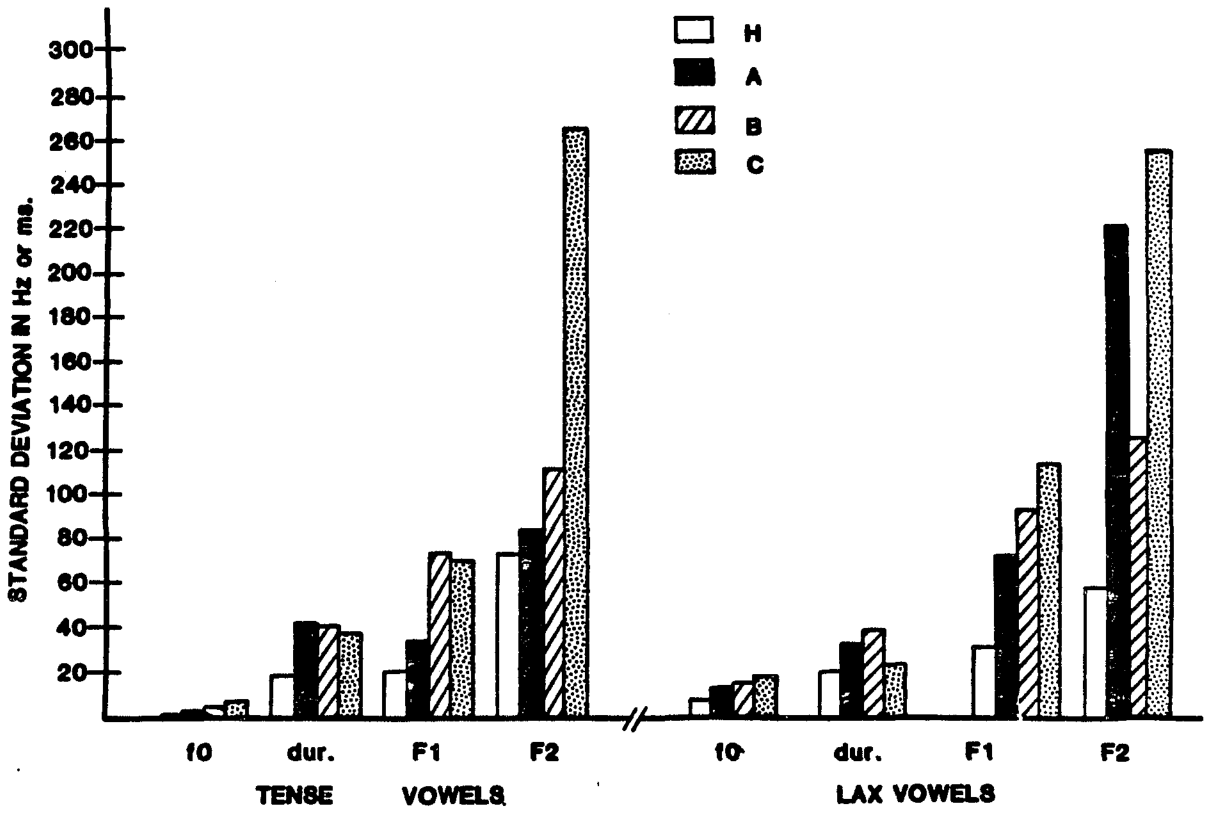

Figure 4.19 Mean Standard Deviations for Groups H, A, B and C for Tense vs. Lax Vowels: f0, Duration, F1, F2.

To determine whether the observed differences were statistically significant, one-way analyses of variance were performed which compared the four subject groups (H, A, B, C) on the basis of their standard deviations for each acoustic parameter. These analyses were performed three times: (1) including all test vowels, (2) including tense vowels only, and (3) including lax vowels only. All three runs indicated significant differences between groups for all acoustic variables. To consider these differences in more detail, post-hoc pairwise comparisons were performed using Tukey's test of honestly significant differences ($p<.05$). The results of these analyses were as follows:

DURATION:

All vowels: All hearing-impaired groups were significantly more variable than the normal controls. No significant differences were found among hearing-impaired groups (A,B,C).

Tense vowels: No differences were found between any of the test groups.

Lax vowels: Groups A and B were significantly more variable than the normal controls. Group C was not.

FUNDAMENTAL FREQUENCY:

All vowels: All hearing-impaired groups were significantly more variable than the normal controls. Group A was significantly less variable than Group C.

Tense vowels: The controls and Group A were significantly less variable than Group C.

Lax Vowels: The only significant differences here were between the controls and Group C.

FIRST FORMANT:

All vowels: Results confirm that all hearing-impaired groups were significantly more variable than the normal controls regarding F1 frequency. Groups B and C were significantly more variable than Group A.

Tense Vowels: There was no significant difference between Group A and the controls. Significant differences were found between Groups B and C, and the normal controls.

Lax Vowels: All hearing-impaired groups were significantly more variable than the controls. Group A was significantly less variable than Group C.

SECOND FORMANT:

All vowels: AIl 3 hearing-impaired groups were more variable than the normal controls. Group C was significantly more variable than Groups A and B.

Tense Vowels: The only significant differences found here were between Group C and the other three groups. That is, no significant differences were found between the normal controls and Groups A and B regarding F2 variability for the point vowels.

Lax Vowels: All hearing-impaired groups were significantly more variable than the controls. Groups A and B were significantly less variable than Group C.

In general, one can conclude the following:

1. When one considers the seven test vowels together, all hearing-impaired groups are significantly more variable than the controls.
2. When one considers a subset of the test vowels, some hearing-impaired groups are not significantly more variable that the controls in the production of certain acoustic features.

Before one can conclude from this that the measured acoustic features were being used to specify vowel identity, one must establish the extent to which they differentiate between vowel targets.

Vowel Distinctiveness:

The phonetic value of F1 and F2 was discussed with reference to the mean and token vowel plots. To consider the phonetic value of f0 and duration, group means as a function of vowel target were plotted and are shown in Figures 4.20 and 4.21. The normalized values of these measures are shown in Appendices C1 and C2 [10]. In a general way, it appears that the shape of the functions are similar for the normal hearing and hearing-impaired groups. The greatest deviations can be seen in f0 for Group C. Here, large increases in f0

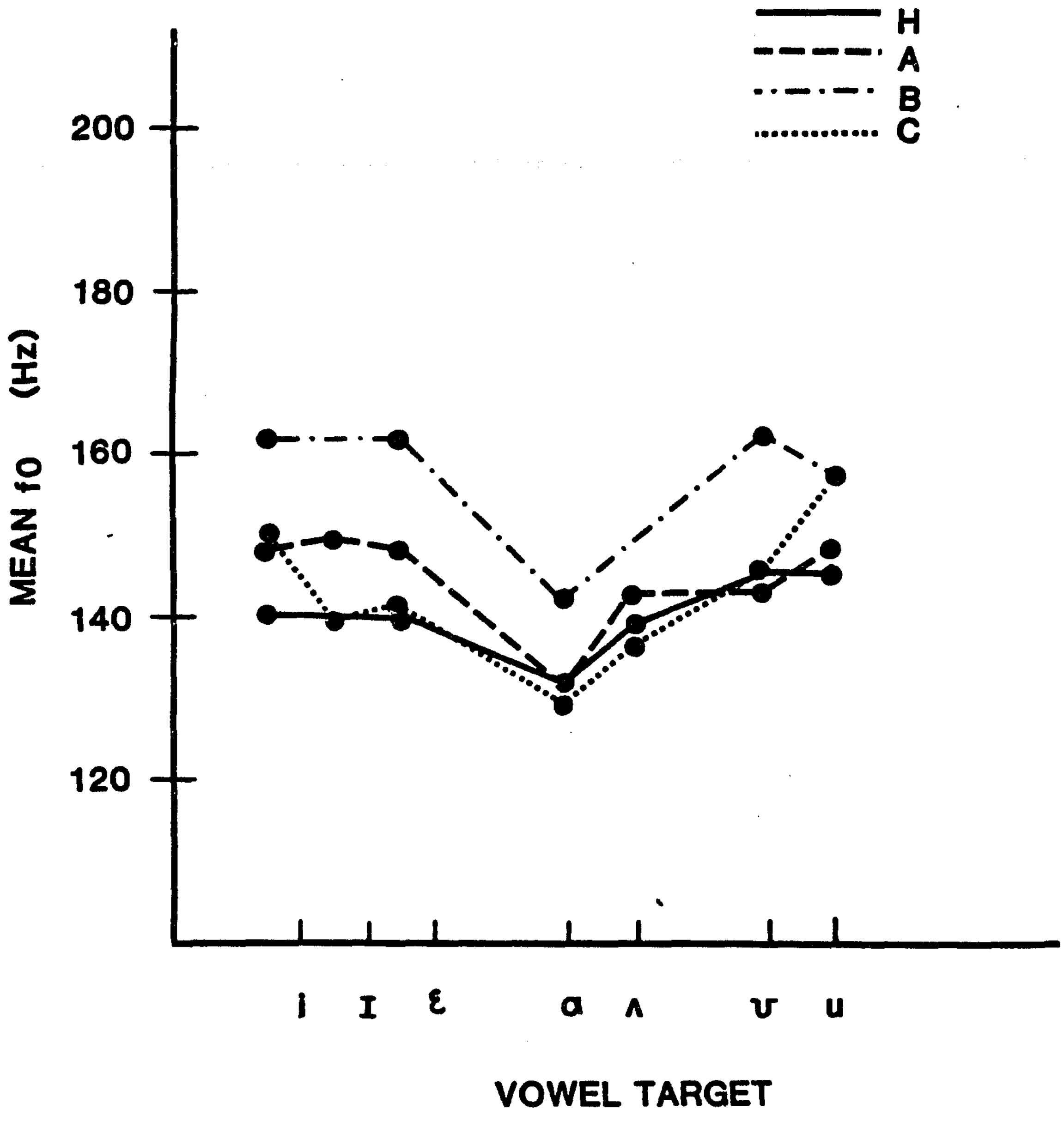

Figure 4.20 Group Means (H, A, B, C) for f0 as a Function of Vowel Target.

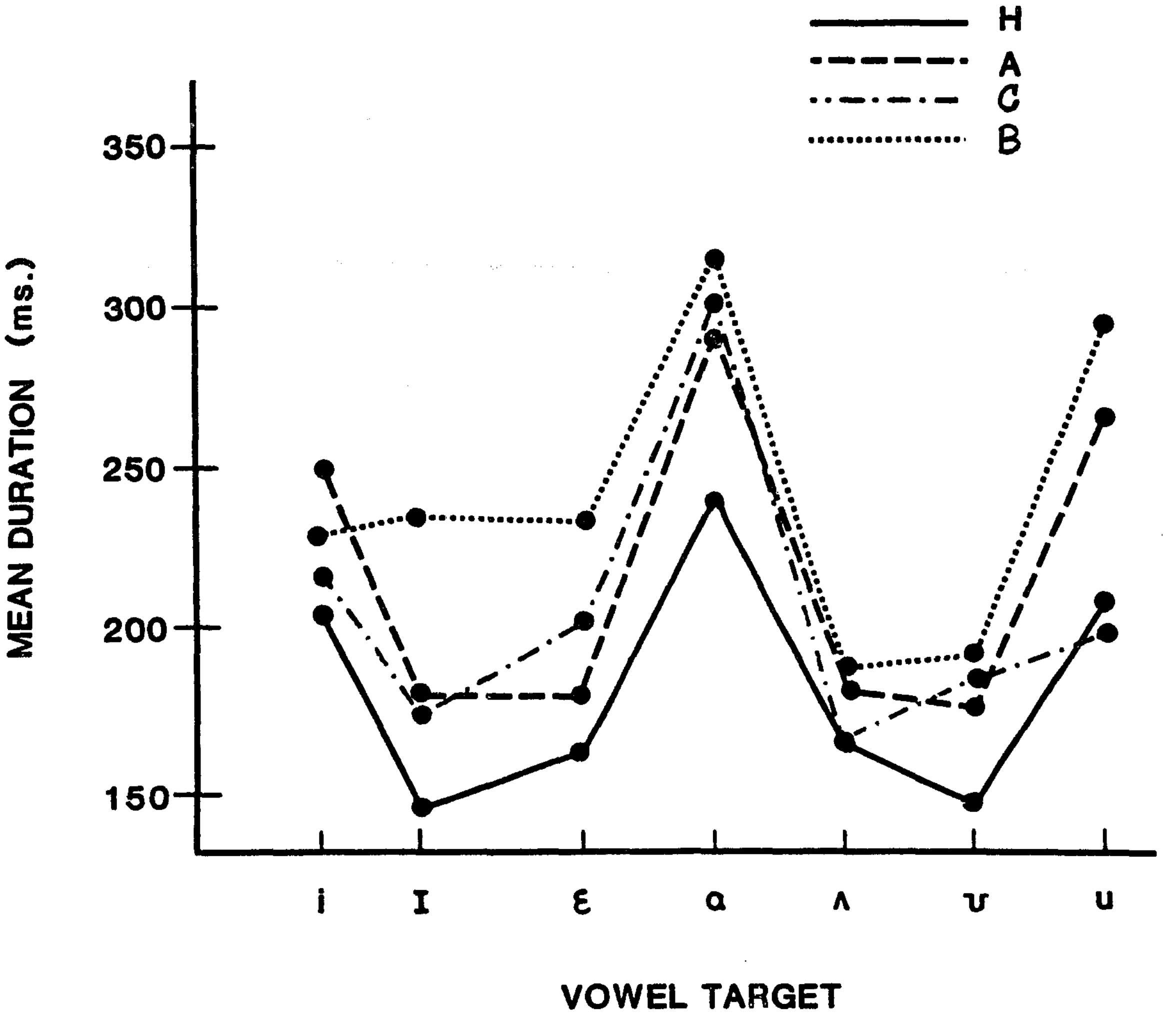

Figure 4.21 Group Means (H, A, B, C) for Duration as a Function of Vowel Target.

for /i/ and /u/ can be noted. Regarding duration, somewhat larger changes between /ʊ/ and /u/ appear for Groups A and B than for the controls.

It was possible that hearing-impaired talkers were differentiating between vowel targets on the basis of parameters, or a combination of parameters, not previously considered. To test vowel distinctiveness statistically, a Linear Discriminant analysis was performed. The analysis was performed twice. On the first run, only the first and second formant frequencies were considered. A second run was performed in which no specifications were used other than that all four acoustic dimensions (f0,F1,F2,duration) enter into the analysis. This provided a direct measure as to whether the addition of alternative parameters and/or a combination of parameters served to reveal a 'deviant but well-defined phonology'.

The order of entry into the analysis is determined by each variable's F-value resulting from a one-way analysis of variance, the independent variable being vowel target. The acoustic parameter with the largest F-value is the most powerful discriminator of vowel type and so is entered first into the discriminant function. In step 2 of the analysis, F-values are calculated for each remaining variable. These F-values are determined by a one-way analysis of covariance; the covariates being the variables which have already been entered (BMDP, 1983). Taking intrasubject variability into account, this step-like procedure isolates the most powerful combination of variables for vowel differentiation for each talker.

Malhalanobis D^2 statistics (Malhalanobis, 1936) were calculated for each token. The D^2 statistic represents the distance, in a multidimensional space, between each token and the vowel target mean. It is in some sense an estimate of 'category goodness'. These measures are used to compute F-values for each target which can be applied, pair-by-pair, to test for the equality of group means. This provides a direct measure of the number of vowel targets that are statistically distinctive on the basis of F1 and F2 values and/or any combination of the four acoustic variables. The results of these paired comparisons should be viewed conservatively since the relationship between statistically significant differences between vowel targets and perceptually significant differences is unknown. Even when the comparisons for the deaf talkers were 'statistically significant', the magnitude of the values in the F-matrices were substantially smaller than any seen for the normal hearing talkers. For this reason, a .001 significance level was adopted.

The results of the two analysis runs for each talker are shown in Table 4.11. Vowels on different lines are significantly different and vowels on the same line are not. Also shown is the 'order of entry' of the variables for each talker.

It can readily be seen that for the most part, vowel targets are no more well specified by the inclusion of f0 and duration information than when discrimination is based on F1 and F2 alone. The exceptions to this were for the comparisons between /ʊ/ and /u/ for talker D1, between /i/ and /I/, and /ʊ/ and /u/ for talker D3, and between /ʊ/ and /u/, and /ɑ/ and /ʌ/, and /I/ and /ɛ/ for talker D6. A one-way analysis of variance was performed to determine which additional variables were associated

Table 4.11

Statistically Discriminable Vowel Targets

TALKER	F2 AND F1	ALL MEASURES	ORDER OF ENTRY
D1	i-I-ɛ	i-I-ɛ	F2
	ɑ	ɑ	F1
	ʌ	ʌ	DUR
	ʊ-u	ʊ	f0
		u	
D2	i-I-ɛ	i-I-ɛ	F2
	ɑ	ɑ	F1
	ʌ	ʌ	DUR
	ʊ-u	ʊ-u	f0
D3	i-I-ɛ	i	F2
	ɑ	I-ɛ	F1
	ʌ-ʊ	ɑ	DUR
	ʊ-u	ʌ-ʊ	f0
		u	

Table 4.11 (cont'd)

Statistically Discriminable Vowel Targets

TALKER	F2 AND F1	ALL MEASURES	ORDER OF ENTRY
D4	i-u	i-u	F1
	I-ε	I-ε	DUR
	ɑ	ɑ	f0
	ʌ-ʊ	ʌ-ʊ	F2
D5	i-ε	i-ε	F1
	I	I	F2
	ɑ	ɑ	DUR
	ʌ	ʌ	f0
	ʊ-u	ʊ-u	
D6	i	i	DUR
	I-ε	I-ʌ-ʊ	F2
	I-ʌ-ʊ	ε	f0
	ɑ-ʌ	ɑ	F1
	ʊ-u	u	

with the clarification of these pairs. For each talker, post-hoc comparisons using Tukey's test of honestly significant differences (p<.05) showed tense/lax paired comparisons to be significant only for duration. For talker D6 however, both f0 and duration distinguished between /ʊ/ and /u/, /ɑ/ vs. /ʌ/ was specified by duration only, and /I/ vs. /ɛ/ was distinguished solely on the basis of f0. These results were substantiated by additional Linear Discriminant analyses which included only F1, F2, and the suggested relevant parameters.

The two normal hearing talkers are not shown since all 7 vowels were significantly different for both runs. For these talkers and for D1-D3, the order of entry of the variables was : (1)F2 (2)F1 (3)duration (4)f0. Talker D5 was similar although F1 was entered first. Talkers D4 and D6 showed unusual hierarchies. For D4, F1 and duration were the most powerful delineaters of vowel targets. For D6, duration and F2 served this role.

Summary of Acoustic Analysis:

The results of an acoustic analysis of the data suggest that the hearing-impaired talkers are more variable than the controls in all measured aspects of vowel production. With the exception of certain vowel pairs, it does not appear that we underestimate an individual's vowel production skills when our analyses are restricted to information about F1 and F2 mean values and the token-to-token variability in these measures.

PERCEPTION OF VOWELS PRODUCED BY HEARING AND HEARING-IMPAIRED TALKERS

Overview:

The primary purpose of this experiment was to determine the presence or absence of calibration and context effects in the perception of vowels spoken by the hearing impaired. A specific interest was to compare the perception of vowels spoken by hearing vs. hearing- impaired talkers. An additional concern was the effect of listener experience. This variable was evaluated on two levels; first, to determine the significance of an overall experienced listener advantage in vowel perception; and second, to estimate the interaction between an experienced listener advantage and the stimulus condition (sentence, word, and gated vowel). As discussed earlier, listeners were required to make a non-phonetic as well as a phonetic judgement. Their ability to identify the hearing status of the talker was analyzed as a function of stimulus condition and listener experience.

The Effects of Stimulus Condition on Vowel Intelligibility:

The six hearing-impaired talkers were treated as three groups of two in accordance with the findings of the acoustic analysis. Since their production styles and overall intelligibility suggested these pairings, it seemed reasonable to evaluate the perception of their utterances in this way. Additionally, this breakdown allowed for an analysis of the relationship between overall intelligibility and the other independent variables discussed.

Percent correct scores were tabulated for each talker and submitted to a 4- way analysis of variance with the factors: Group (H, A, B, C), Vowel (/i,I,ɛ,ɑ,ʌ,ʊ,u/), Condition (sentence, word, gated vowel), and Listener (experienced, inexperienced). Percent correct scores were transformed into arcsine units to stabilize the error variance (Brownlee, 1965). While these data were used for this analysis, the figures and tables shown here reflect actual percent correct scores. The results of the analysis are shown in Table 4.12. The name of each factor is abbreviated as follows: Group=G, Vowel=V, Condition=C, and Listener=L. Significant main effects were found for Group, Vowel, and Condition ($p < .01$). Significant interactions were found between Vowel and Condition, and Vowel and Listener. Additionally, the three-way interaction between Vowel, Condition and Listener was significant.

As expected, the four subject groups were not equivalent with respect to overall vowel intelligibility. Average percent correct vowel recognition for the 4 groups of talkers is shown in Table 4.13 along with

Table 4.12

Four Way Analysis of Variance: All Conditions

	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROBABILITY
Group	22.322	3	7.441	33.09	0.0028*
Error	0.899	4	0.225		
Vowel	21.496	6	3.583	11.73	0.0000*
VG	12.562	18	0.698	2.29	0.0298
Error	7.328	24	0.305		
Condition	1.152	2	0.576	28.46	0.0002*
CG	0.159	6	0.027	1.31	0.3521
Error	0.162	8	0.020		
VC	1.592	12	0.133	6.65	0.0000*
VCG	0.867	36	0.024	1.21	0.2683
Error	0.957	48	0.020		
Listener	0.003	1	0.003	1.23	0.3289
LG	0.015	3	0.005	2.33	0.2158
Error	0.009	4	0.002		
VL	0.079	6	0.013	3.66	0.0100*
VLG	0.070	18	0.006	1.53	0.1630
Error	0.087	24	0.004		
CL	0.011	2	0.005	2.97	0.1087
CLG	0.008	6	0.001	0.69	0.6635
Error	0.015	8	0.002		
VCL	0.101	12	0.008	2.86	0.0049*
VCLG	0.053	36	0.001	0.50	0.9845
Error	0.142	48	0.003		

*significant at $p < .01$

Table 4.13

Percent Correct Vowel Recognition for Individuals and Groups

(mean of the three conditions)

Talker	% Correct	Group	% Correct
D1	55.0	Group A	55.5
D2	47.5	Group B	51.1
D3	54.3	Group C	46.6
D4	55.7	Group H	87.3
D5	40.0		
D6	53.8		
H1	88.5		
H2	80.0		

mean percent correct scores for individual talkers. Group scores ranged from 87.3% correct for the normals to 46.6% correct for Group C. Individual scores ranged from 88.5% to 40.0% correct.

As mentioned, the main effect of Vowel was significant. The intelligibility of each test vowel is shown in Table 4.14 as a function of Group. It can be seen that the relative intelligibility of vowels changes little as a function of the hearing- impaired groups. The vowel /ɑ/ is, in all cases, the most intelligible vowel and /ɛ/ , the least intelligible. The other point vowels, /i/ and /u/ are second and third in percent correct identification for Groups B and C. For Group A, /I/ is the second most intelligible vowel, followed by /i/ and then /u/. It appears that while the controls show a different order, variability was such that this difference was not statistically significant.

The interaction between Listener and Vowel was also significant. Mean percent correct scores for experienced and inexperienced listeners as a function of vowel target are shown in Table 4.15. Experienced listeners showed slightly superior scores for all vowels except /i/. On the whole, however, these differences were small, and the main effect of listener experience was not significant.

Table 4.16 shows the percent correct scores for experienced and inexperienced listeners as a function of Group. There was no significant difference between the two groups of listeners in any of the talker groups. Additionally, there was no interaction between Condition and Listener. The equivalence of the two groups of listeners regarding overall vowel intelligibility scores as a function of Condition is shown graphically in Figure 4.22 for the hearing and hearing-impaired talkers.

Table 4.14

Percent Correct Vowel Recognition as a Function of Vowel Target

(mean of the 3 conditions)

	H	A	B	C	Mean HI
/i/	96.0	60.5	63.0	67.0	63.2
/I/	96.1	84.9	44.1	20.4	59.2
/ɛ/	93.8	15.7	21.5	12.3	17.3
/ɑ/	84.4	89.2	90.4	82.7	87.4
/ʌ/	85.2	35.8	24.3	23.9	28.5
/ʊ/	32.9	30.8	28.2	14.3	26.3
/u/	94.3	48.9	66.0	34.3	51.3

Table 4.15

Percent Correct Vowel Recognition as a Function of Vowel Target:
Experienced vs. Inexperienced Listeners

	/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/
Experienced	70.3	70.1	37.5	87.1	45.5	29.2	64.6
Inexperienced	74.0	67.8	37.0	86.3	46.4	26.1	60.5
Mean	72.2	69.0	37.3	86.7	45.9	27.7	63.1

Table 4.16

Percent Correct Vowel Recognition as a Function of Group:

Experienced vs. Inexperienced Listeners

	H	A	B	C	Mean (all Groups)
Experienced	87.3	55.9	51.8	46.0	61.1
Inexperienced	87.3	54.3	50.4	47.2	60.5

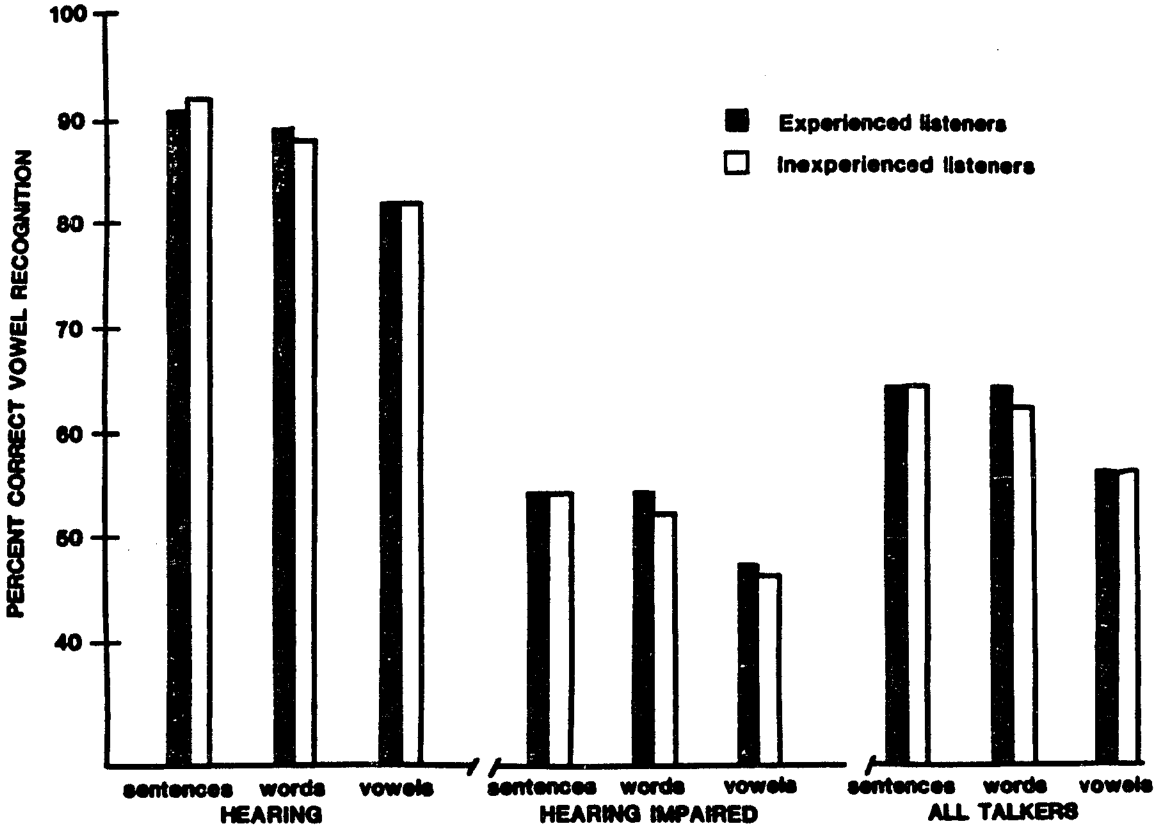

Figure 4.22 Vowel Recognition as a Function of Condition: Experienced vs. Inexperienced Listeners.

The main effect of Condition was significant (p<.01) as were its interactions with the following factors: Vowel x Condition, and Vowel x Listener x Condition. Scores by Vowel and Condition for individual talkers are reported in Appendices D-F. These results averaged over hearing and hearing-impaired talkers are shown in Appendix G. Table 4.17 shows mean percent correct vowel identification as a function of stimulus condition for individual talkers. These values for the test groups are reported in Table 4.18 and are shown graphically in Figure 4.23. These data will be discussed in more detail below.

Calibration and Context Effects:

In order to investigate more fully the main effects of Condition and its interactions with other test variables, two separate analyses of variance were performed. The first compared only the Sentence and Word conditions. Results from this analysis refer directly to the calibration effect. The second compared only the Word and Gated Vowel conditions and provides information regarding the consonant context effect.

Table 4.17

Percent Correct Vowel Recognition as a Function of Condition: Individual Talkers

Talkers	Sentences	Words	Gated Vowels
D1	57.9	55.2	51.8
D2	49.9	50.2	42.2
D3	59.8	55.2	47.9
D4	56.9	58.6	51.7
D5	41.1	41.2	37.1
D6	58.0	57.5	45.8
H1	93.8	89.6	82.2
H2	88.9	88.1	81.0

Table 4.18

Percent Correct Vowel Recognition as a Function of Condition: Group

	H	A	B	C
Sentences	91.4	58.2	53.8	49.3
Words	88 8	57.0	52.6	49.2
Gated Vowels	81.6	50.0	46.9	41.3
Mean	87.3	55.1	51.1	46.6

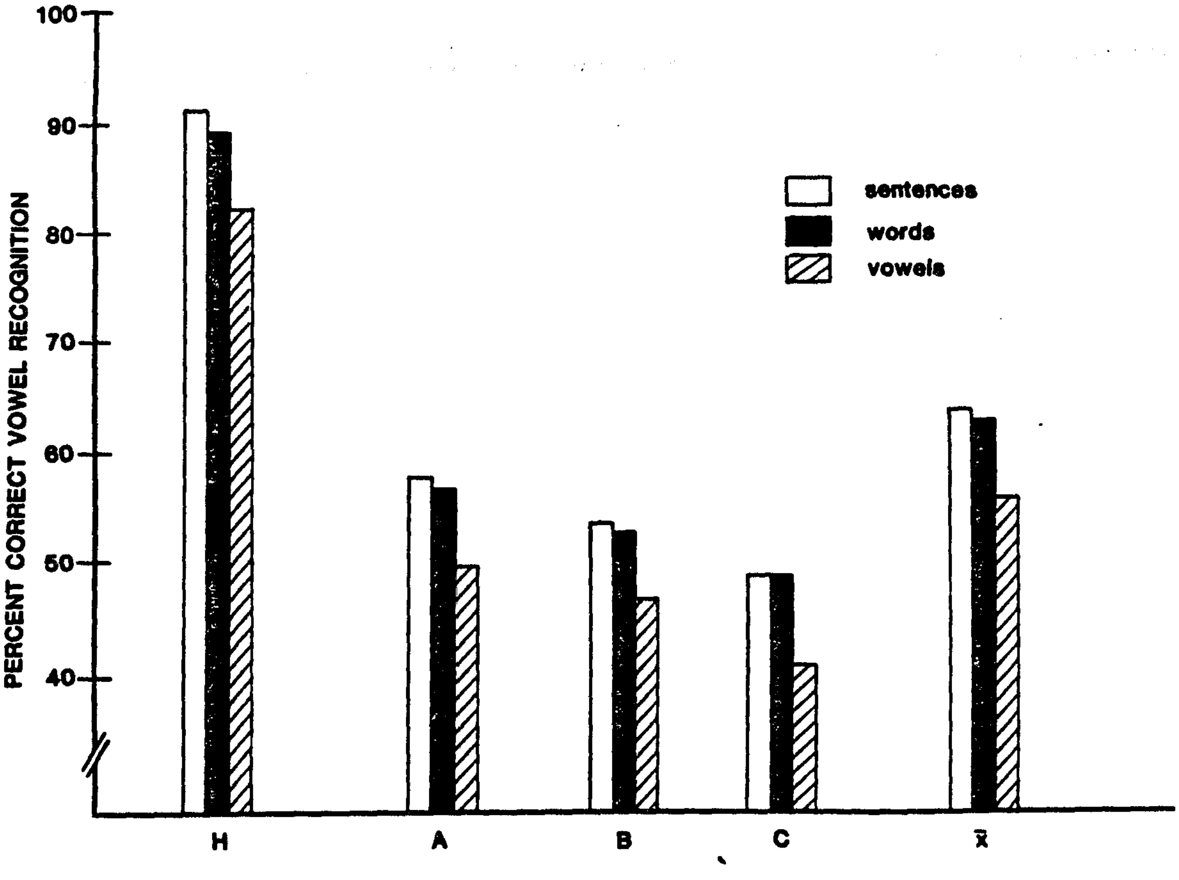

Figure 4.23 Vowel Recognition as a Function of Condition and Group.

Calibration Effect:

A 4-way analysis of variance was performed with the factors: Group, Vowel, Condition (sentence and word) and Listener. The results of this analysis are shown in Table 4.19. As expected based on the previous analysis, the effects of listener experience were not significant. Again, the effects of Group were significant. Mean percent correct scores averaged across the two conditions were 90.1%, 57.6%, 53.2% and 49.3% for groups H, A, B, and C respectively.

The main effect of Vowel was also significant as were the following interactions: Vowel x Group, Vowel x Listener, and Vowel x Condition x Listener.

Vowel x Group:

Tables 4.20 and 4.21 show the percent correct identification for individual vowels as a function of Group for the Sentence and Word conditions, respectively. As discussed earlier, the hearing-impaired groups all showed similar patterns regarding the relative intelligibility of vowels. Presumably, the Vowel x Group interaction reached significance here because the normal hearing talkers showed a consistently different pattern than that of the hearing-impaired talkers. For normal talkers, /I,u,i/ were the most intelligible vowels followed by /ɛ,ʌ,ɑ/. /ʊ/ was the least intelligible vowel for this group. The rank order of vowels with respect to percent correct identification is shown below for the normal hearing and hearing-impaired talkers in the two test conditions:

Sentences		Words	
HI	H	HI	H
ɑ	I	ɑ	I
I	u	i	u
u	i	I	i
i	ɛ	u	ʌ
ʌ	ʌ	ʌ	ɛ
ʊ	ɑ	ʊ	ɑ
ɛ	ʊ	ɛ	ʊ

Note that for both groups, apart from one or two minor reversals, the rank order is the same for both stimulus conditions.

Vowel x Listener:

The interaction between Vowel and Listener was significant. However, the differences between the groups was never more than 3%; the predicted advantage reversing for the vowels /i/ and /ʌ/.

Table 4.19

Four Way analysis of variance: calibration effect

	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.
Group	16.371	3	5.441	27.50	0.0039*
Error	0.793	4	0.198		
Vowel	17.818	6	2.970	15.92	0.0000*
VG	9.733	18	0.541	2.90	0.0079*
Error	4.476	24	0.187		
Condition	0.038	1	0.039	2.30	0.2036
CG	0.040	3	0.014	0.81	0.5498
Error	0.067	4	0.017		
VC	0.379	6	0.063	7.88	0.0001*
VCG	0.231	18	0.013	1.60	0.1406
Error	0.192	24	0.008		
Listener	0.002	1	0.002	0.72	0.4428
LG	0.017	3	0.006	2.08	0.2461
Error	0.011	4	0.003		
VL	0.123	6	0.020	4.18	0.0051*
VLG	0.096	18	0.005	0.98	0.5126
Error	0.118	24	0.005		
CL	0.011	1	0.011	5.95	0.0713
CLG	0.004	3	0.001	0.77	0.5671
Error	0.007	4	0.002		
VCL	0.051	6	0.008	3.79	0.0085*
VCLG	0.029	18	0.002	0.72	0.7563
Error	0.053	24	0.002		

*significant a p < .01

Table 4.20

Percent Correct Vowel Recognition as a Function of Vowel:

Sentence Condition

	H	A	B	C
/i/	97.2	61.7	57.3	61.1
/I/	99.7	94.1	58.1	30.0
/ɛ/	93.5	8.4	15.9	11.6
/ɑ/	92.3	93.3	94.3	86.9
/ʌ/	93.0	38.8	20.0	24.6
/ʊ/	40.4	30.3	18.9	17.0
/u/	97.4	58.8	83.9	47.0

Table 4.21

Percent Correct Vowel Recognition as a Function of Vowel:

Word Condition

	H	A	B	C
/i/	95.2	60.3	72.9	73.6
/I/	97.4	91.1	39.6	22.5
/ɛ/	93.9	18.0	21.3	11.9
/ɑ/	86.1	95.2	94.8	87.4
/ʌ/	94.0	42.4	22.5	27.9
/ʊ/	31.7	30.5	26.9	10.5
/u/	95.7	44.0	68.2	34.0

Vowel x Condition x Listener:

Figure 4.24 shows the difference scores (% correct Sentence condition - % correct Word condition) obtained as a function of vowel target for experienced vs. inexperienced listeners. Points below the 'zero line' indicate superior performance for the Word over the Sentence condition. For /I,ɑ,u,ʊ/, the Sentence condition yielded higher intelligibility scores than the Word condition. For /i,ɛ,ʌ/, however, the predicted effect was reversed for both groups of listeners. Difference scores ranged from -6% to +10%. With regards to the 3-way interaction, a consistent trend may be noted. The slope of the lines in Figure 4.24 is positive for all tense vowels and negative or zero for three out of four of the lax vowels. That is, the two groups of listeners seem to respond differently to tense vs. lax vowels as a function of stimulus condition. In a relative way, inexperienced listeners benefit more than experienced listeners in tense vowel recognition when provided with a precursor sentence. Alternatively, experienced listeners show more improvement than do inexperienced listeners in the recognition of lax vowels as a result of the precursor.

The main effect of Condition was not significant here. Neither were the interactions between Condition and Listener, nor Condition and Group. That is, there was no statistical evidence of superior performance in vowel recognition when listeners were or were not provided with exemplars of the talker's point vowels. The equality of the two stimulus conditions was consistent across all groups of talkers and for both groups of listeners.

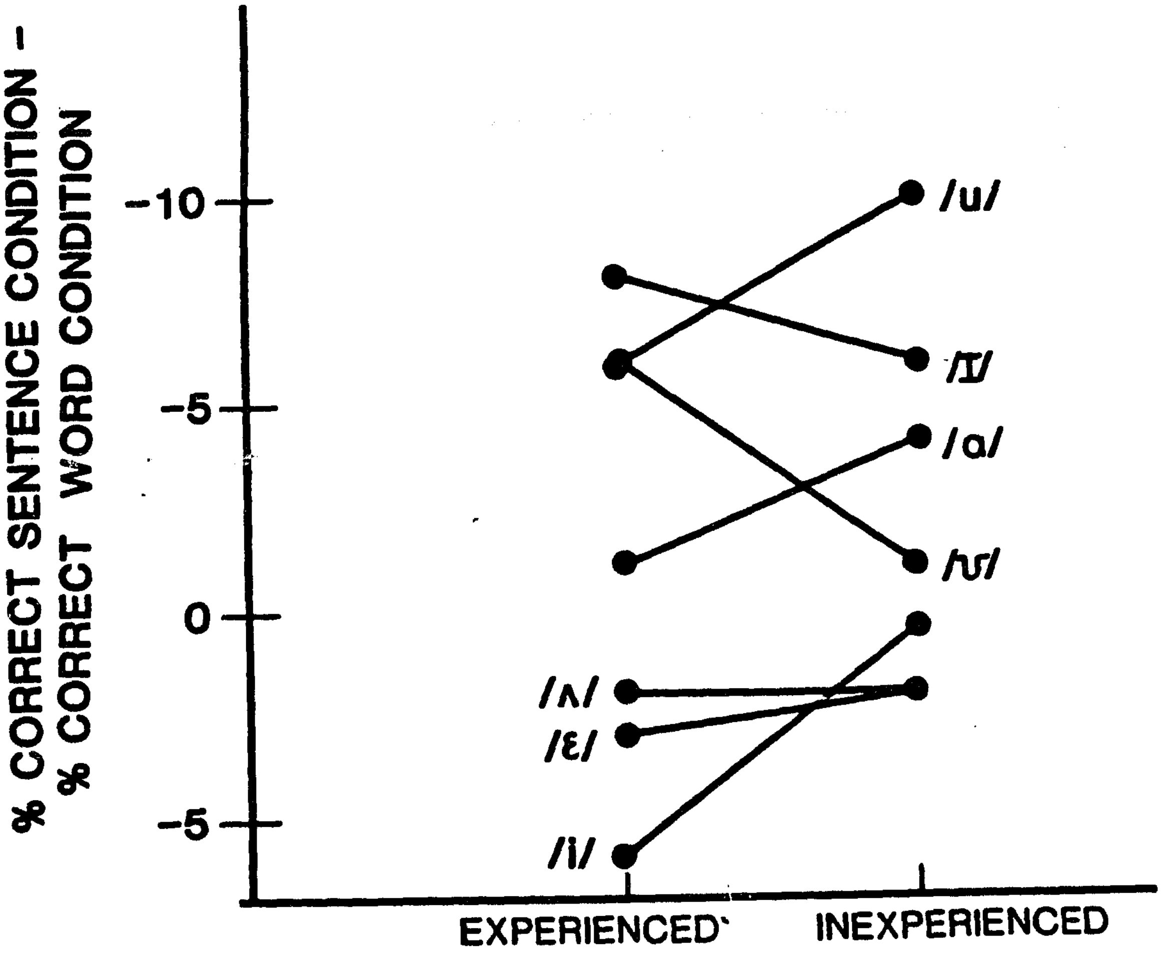

Figure 4.24 Difference Scores Between Sentence and Word Condition as a Function of Vowel and Listener.

Consonant Context Effect:

A 4-way analysis of variance was performed with the factors: Group, Vowel, Condition (word and gated vowel) and Listener. The results of the analysis are shown in Table 4.22.

The main effects of Group and Vowel were significant. The mean percent correct scores across the Word and Gated Vowel conditions were 85.2%, 53.5%, 49.8% and 45.3% for groups H, A, B, and C, respectively. Tables 4.21 and 4.23 show the intelligibility of individual vowels by Group for the Word and Gated Vowel conditions, respectively. Shown below are their rank orderings for hearing vs. hearing-impaired talkers:

Words		Gated Vowels	
HI	H	HI	H
ɑ	I	ɑ	i
i	u	i	ɛ
I	i	I	I
u	ʌ	u	u
ʌ	ɛ	ɝ	ɑ
ɝ	ɑ	ʌ	ʌ
ɛ	ɝ	ɛ	ɝ

The relative intelligibility of vowels, as discussed earlier, changes little across conditions. One exception to this was for the hearing talkers.

Table 4.22

Four Way Analysis of Variance: Context Effect

	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROBABILITY
Group	13.444	3	4.481	31.46	0.0031*
Error	0.570	4	0.142		
Vowel	12.781	6	2.130	9.11	0.0000*
VG	8.271	18	0.450	1.97	0.0611
Error	5.610	24	0.234		
Condition	0.665	1	0.665	33.39	0.0045*
CG	0.042	3	0.014	0.70	0.6015
Error	0.080	4	0.020		
VC	0.637	6	0.106	5.14	0.0016*
VCG	0.564	18	0.031	1.52	0.1678
Error	0.496	24	0.021		
Listener	0.009	1	0.009	12.58	0.0239
LG	0.004	3	0.001	1.95	0.2642
Error	0.003	4	0.001		
VL	0.036	6	0.006	1.92	0.1188
VLG	0.085	18	0.005	1.49	0.1800
Error	0.076	24	0.003		
CL	0.003	1	0.003	2.09	0.2215
CLG	0.085	3	0.000	0.11	0.1800
Error	0.076	4	0.001		
VCL	0.028	6	0.005	2.46	0.0536
VCLG	0.030	18	0.002	0.89	0.5992
Error	0.045	24	0.002		

*significant at $p < .01$

Table 4.23

Percent Correct Vowel Recognition as a Function of Vowel:

Gated Vowel Condition

	H	A	B	C
/i/	95.5	59.3	58.8	66.1
/I/	91.3	69.5	34.4	8.8
/ɛ/	93.9	20.7	27.2	13.4
/ɑ/	74.8	79.1	82.1	73.8
/ʌ/	68.5	26.2	30.3	19.2
/ʊ/	26.7	31.6	38.9	15.5
/u/	89.8	43.8	45.9	22.0

In the Gated Vowel condition, it appears that the front vowels are uniformly more intelligible than the back vowels. This occurs as the result of the interaction between Vowel and Condition as will be discussed below.

The main effect of Condition was significant (p<.01). That is, vowels were more intelligible when heard in their original consonant context than when heard as steady state tokens. Figure 4.25 shows the context effect as a function of Group. The difference scores were 7.2%, 7.0%, 5.7%, and 8.0%, for groups H, A, B, and C, respectively. The interactions between Listener and Condition and between Group and Condition were both non-significant, as was the main effect of listener experience.

Vowel x Condition:

The interaction between Vowel and Condition was significant. Figure 4.26 shows the percent correct scores for each vowel as a function of stimulus condition for all talkers. As can be seen, the context effect was minimally reversed for the vowels /ɛ/ and /ʊ/. Interestingly, the context effect was not a function of the overall intelligibility of the vowel.

Figure 4.27 shows the context effect for hearing talkers. The relative intelligiblity of vowels changes for this group in the Gated Vowel condition because /i/ and /ɛ/ do not show a decrement in performance as a function of consonant context as do the other test vowels. Thus, their relative position in the group improves in this condition.

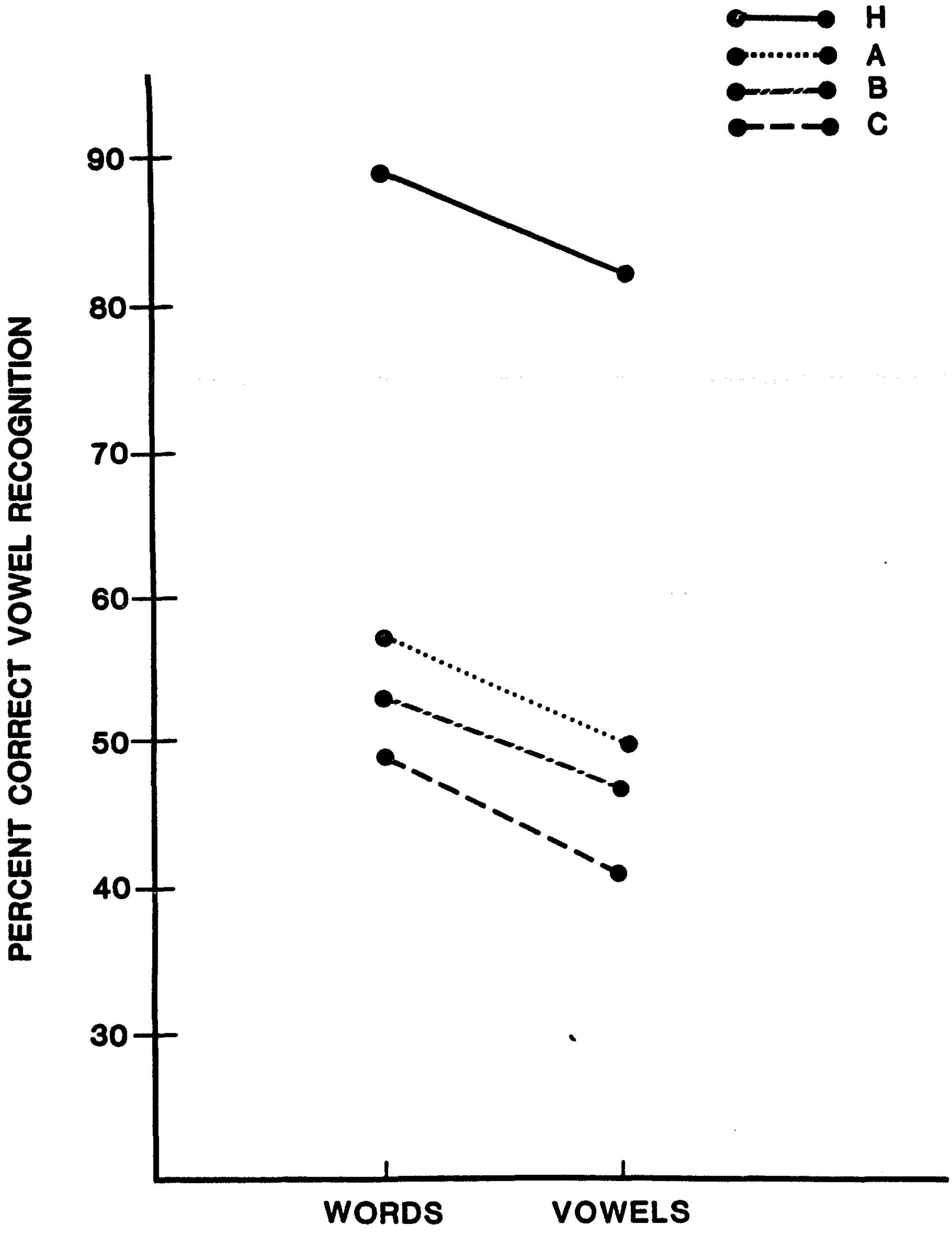

Figure 4.25 Vowel Recognition as a Function of Condition: Context Effect.

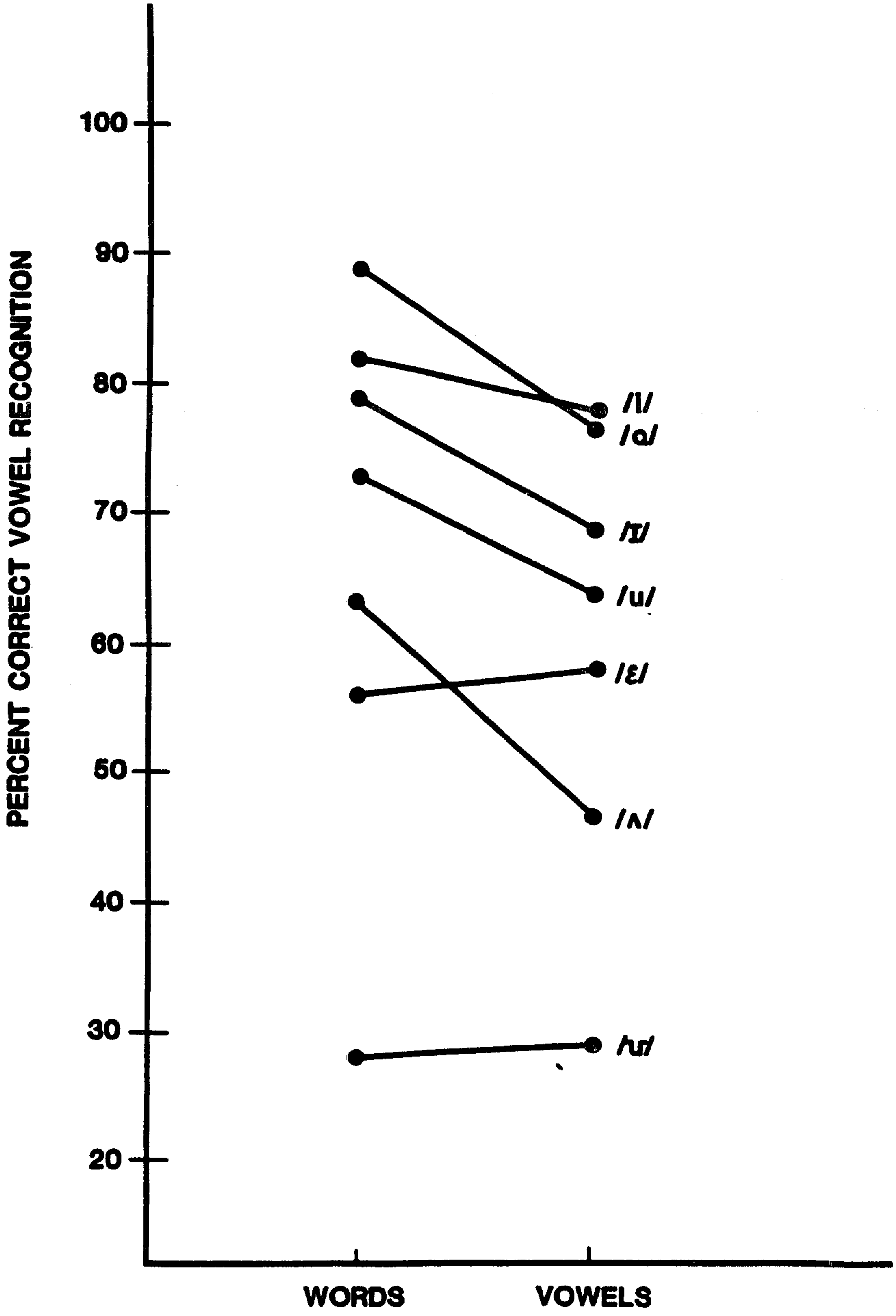

Figure 4.26 Vowel Recognition as a Function of Condition and Vowel: All Talkers.

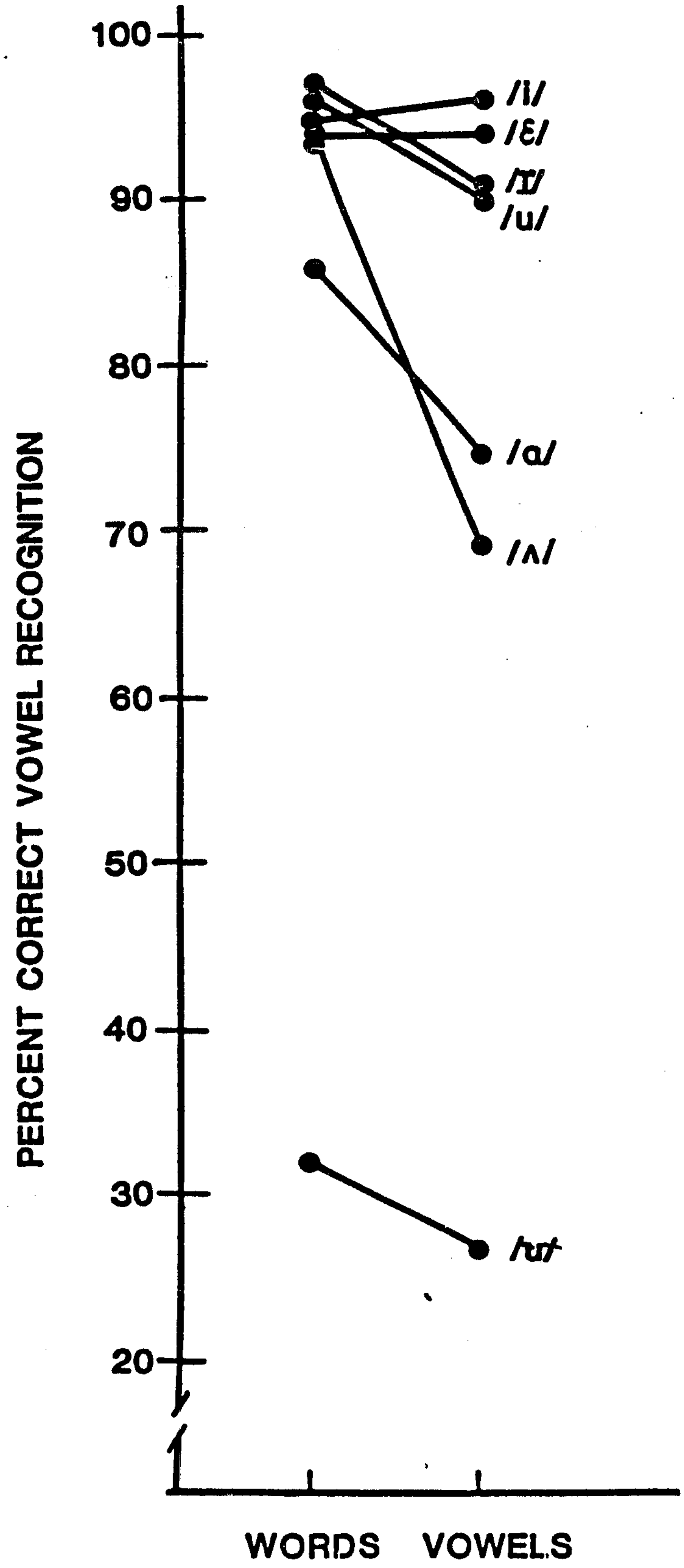

Figure 4.27 Vowel Recognition as a Function of Condition and Vowel: Hearing Talkers.

Summary:

The results of analyses concerning the perception of vowels spoken by hearing and hearing-impaired talkers may be summarized as follows:

1. There were no significant differences between the hearing and hearing-impaired talkers regarding calibration or context effects.
2. The calibration effect was not significant for any group of talkers.
3. The context effect was significant for all groups of talkers.
4. There were no significant differences between experienced and inexperienced listeners in overall vowel recognition scores.
5. Experienced and inexperienced listeners showed the same pattern of results for all conditions and all groups of talkers.

Effects of Condition on Identification of Hearing Status:

A two-way analysis of variance was performed to investigate the effects of Condition on the ability of listeners to identify the hearing status of the talkers. There were 6 hearing-impaired and only 2 normal hearing talkers in the subject pool. These unequal n's were likely to create a response bias that would influence the results both within and across conditions. To circumvent this, a d' score was derived for each listener in each condition. These scores were used to evaluate the significance of the independent variables. The outcome of this analysis is shown in Table 4.24.

Results indicate highly significant effects for Condition ($p<.001$) and no effect of listener experience. d' scores as a function of condition are shown in Figure 4.28 and it is immediately obvious that listeners' ability to identify the talker as normal hearing vs. hearing-impaired decreases as articulatory information over time is removed.

As discussed in Chapter II, Calvert's (1961) results suggested that listeners cannot judge the hearing status of the talker in the absence of dynamic articulatory cues. A specific interest then, was listeners' ability to perform this task in the Gated Vowel condition. A one-sample t-test was performed to test whether or not d' scores were significantly above chance in this condition ($d'>0$). The null hypothesis proved to be false ($p<.01$). That is, both groups of listeners performed significantly above chance in identifying the hearing status of the

Table 4.24

Analysis of Variance:

Identification of Hearing Status of the Talker

SOURCE	SUM OF SQUARES	DEGREES OF FREEDOM	MEAN SQUARE	F	TAIL PROB.
Listener	0.403	1	0.403	0.60	0.4408
error	65.871	98	0.672		
Condition	536.393	2	268.20	907.31	0.0000*
CL	1.673	2	0.837	2.83	0.0614
error	57.937	196	0.296		

*significant at $p < .001$

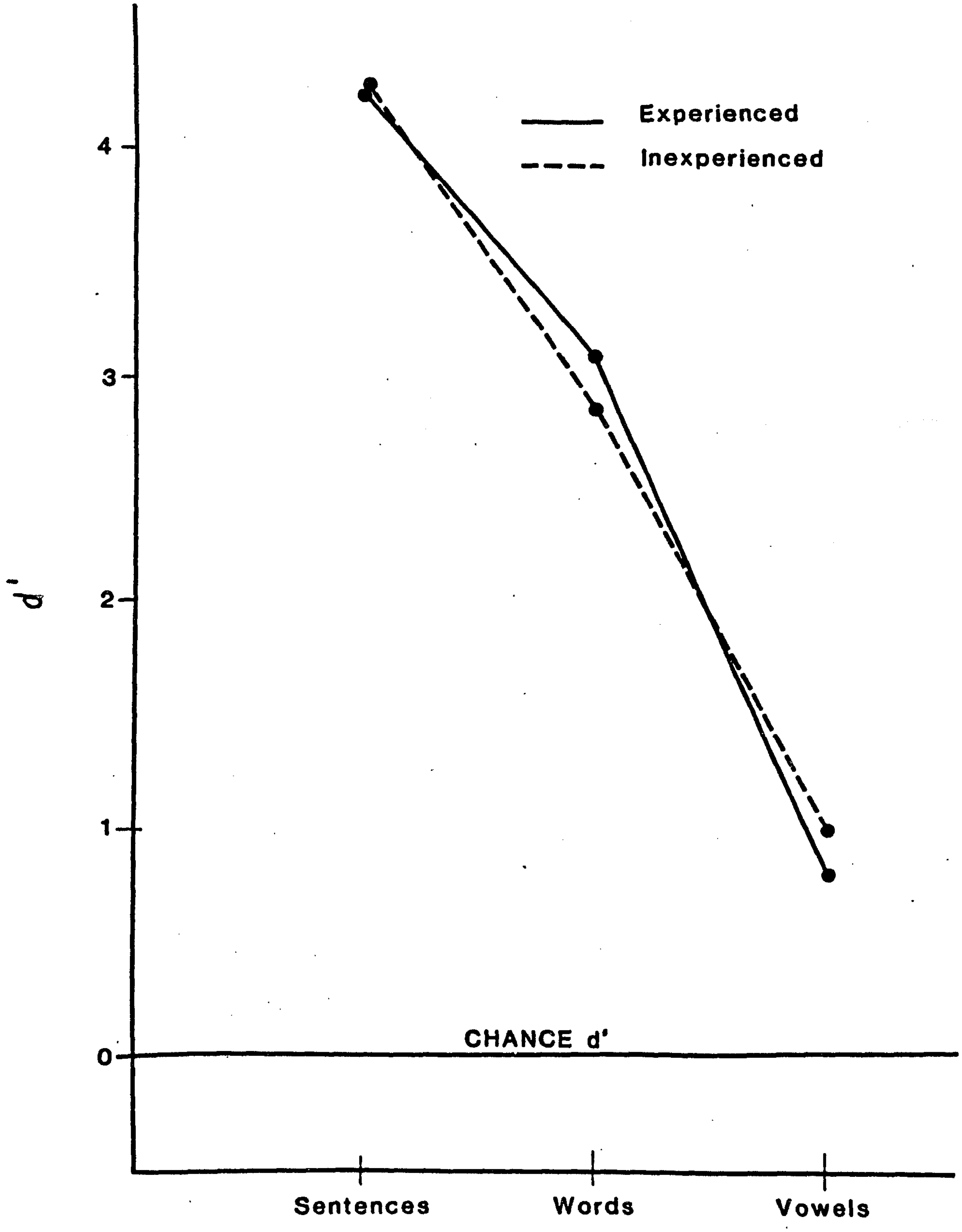

Figure 4.28 Recognition of Hearing Status as a Function of Condition: Experienced vs. Inexperienced Listeners.

talker given only steady state information.

Perceptual Criteria:

Listeners in this study were asked to perform two perceptual tasks simultaneously. One task was phonetic (vowel identity) and the other non-phonetic (hearing status). An interesting question is whether these two perceptual decisions are dependent on one another in some way. That is, do listeners correctly identify an utterance as a given vowel as a result of correctly identifying the talker as hearing impaired? Alternatively, is their correct identification of the talker as hearing impaired dependent in some way on their success at the phonetic task? A 2x2 chi square analysis was performed to test the independence of these two perceptual tasks.

Each hearing-impaired talker was analyzed individually as a function of stimulus condition and listener experience (36 chi square tests). The majority of chi squares (21 out of 36) were not significant at the .05 level. That is, the relationship between the phonetic and non-phonetic task was not significant. Using a .001 level of significance, an obvious pattern emerged. For a single talker (D3) test results were significant for both groups of listeners in the Word and Gated Vowel conditions. For one other talker (D1), chi square results were significant for naive listeners in the Gated Vowel condition. These talker/condition combinations were subjected to additional chi square tests as a function of vowel type.

For talker D3, /ɑ,ʊ,u/ showed significant effects in the Word condition by experienced listeners. /I/ was significant for both groups of listeners in the Gated Vowel condition. /i/ was significant in the Word and Gated Vowel conditions for naive listeners ($p<.05$). For Talker D1, /i/ and /u/ were significant in the Gated Vowel condition by naive listeners ($p<.05$).

For these test situations then, the perceptual tasks were apparently not independent of one another. The raw data revealed the following unexpected relationship: When a listener correctly identifies the talker as deaf, she/he is more likely to *misidentify* the vowel than if she/he incorrectly identifies the speaker as hearing. The notion that listeners use different (and more appropriate) criteria for vowel identification when they know the speaker is deaf would have predicted the opposite relationship.

Analysis of Error Patterns:

Confusion matrices were generated for each talker as a function of Condition and Listener. These matrices were used in a comparative analysis of error patterns for experienced vs. inexperienced listeners. As will be discussed below, the matrices were also used to compare error patterns in the Word vs. Gated Vowel conditions.

Although the overall percent correct vowel recognition scores showed no significant differences between experienced and inexperienced listeners, it was possible that the differences between the two groups rested in the error patterns they generated. To test this hypothesis, two kinds of chi square analyses were performed. The first compared error matrices (diagonal eliminated) of the two groups of listeners on a row-by-row basis. Each talker and condition was considered separately. Pairs of cells were eliminated if either had a zero entry. Results showed no significant effects of listener experience on the error patterns for any talker in any condition ($p<.01$).

Because cells with zero entries were eliminated, this analysis could not adequately explore the possibility that listeners differed with regards to their spread of error responses across vowel type. The second set of chi square analyses, then, were 2x2, and based on a row-by-row comparison of : (1) the number of responses which constituted the major error response and (2) the sum of all other error responses. In 126 chi square tests (6 hearing- impaired talkers x 7 vowels x 3 conditions) only five were significant ($p<.01$). In the Sentence condition, the only significant effect was seen for Talker D3 producing /I/. Here, the expected pattern was reversed. That is, the experienced listeners showed a greater spread of errors than the inexperienced listeners. In the Word condition, there was a significant difference between the groups of listeners for /i/ spoken by Talkers D1, D2, and D5. In the Gated Vowel condition, the 'spread of errors' effect was significant for Talker D5, in response to /ʌ/. These results suggest little difference between listeners either in overall ability or error pattern.

Analyses of variance indicated a significant difference between the Word and Gated Vowel condition in percent correct vowel identification (that is, the context effect was significant). To understand the specific nature of the perceptual effect, the confusion matrices for the Word and Gated Vowel conditions (summed over the two groups of listeners) were compared with reference to their error patterns. The most obvious and consistent trend noted was the decrease in cross-triangle confusions (that is, identifying a front vowel as a back vowel and vice versa) in the Word condition. With the vowels /ɑ/ and /ʌ/ excluded from the count, the total number of cross-triangle confusions across all 8 talkers was determined for the two stimulus conditions. Results revealed a total of 1779 cross-triangle confusions in the Gated Vowel condition and 1180 cross-triangle confusions in the Word condition [11]. The decrease of this error type in the Word vs. Gated Vowel condition would suggest that inappropriate and/or ambiguous steady state cues were being clarified by information as to whether the tongue was moving towards the front or the back of the oral cavity.

RELATIONSHIP BETWEEN ACOUSTICS AND PERCEPTION

Questions still remain as to the criteria used by listeners in determining 'acceptable' tokens of a given vowel target. While statistical probabilities of group membership are interesting from the productive side, it requires a substantial leap of faith to assume that statistically distinct tokens will be perceived with a high degree of accuracy. For example, formant values need not be anything like those of normal hearing speakers in order to be statistically resolvable.

In an attempt to understand the relationship between statistical and perceptual significance in vowel identification, comparisons were made between predicted and actual classification scores. As part of a Linear Disriminant analysis, each token is classified as belonging to one of the test groups (vowel type). Two methods of classification are used. The one reported here is called the 'jacknife method' and does not include data from the token being evaluated in the classification routine. That is, it assumes no a priori knowledge of the token being classified. This more closely resembles the perceptual process. Several classification functions were generated by entering a variety of combinations of acoustic parameters into the discriminant function. A non-parametric correlation coefficient (Spearman rank order correlation) was used to measure the association between the predicted percent correct scores and those perceptually determined. Table 4.25 shows the acoustic variables entered and the resulting correlations between the two classification functions. The row labelled 'best fit' reflects inclusion of the most statistically powerful set of acoustic variables (as determined by the LDA program). Perceptual results from both Word and Gated Vowel conditions were used, as appropriate for the acoustic parameters entered. The highest correlations with perception were for F1 and F2 information alone. That is, when the analysis program used information regarding f0 and duration, it generated classification scores that were less closely related to perceptual findings.

The following qualification should be noted. For a variety of reasons, correspondence between perceptual and statistical classification, with regards to their respective sources of information, was far from complete. For example, if transitional information had been entered into the discriminant function, it might have improved the correlation between predicted and obtained classification scores in the Word condition. Moreover, a separate discriminant function was determined for each speaker (that is, vowel tokens were judged solely with reference to the individual's vowel system). This parallels a blocked speaker design more than the randomized speaker design used here.

Token plots were drawn which included those tokens perceived correctly by at least 70% of the listeners. Since the perceptual patterns generated by experienced and inexperienced listeners were so similar, only the functions for naive listeners are shown here to demonstrate the findings. Figures 4.29-4.36 show the F1 and F2 values of the 'correct' tokens for each talker in each of the three stimulus conditions. Lines have been drawn to enclose the points which represent individual vowel targets.

The most strikingly obvious feature of these plots is the near elimination of overlap between vowel categories. That is, those tokens which occupy mutually exclusive F1/F2 areas in a given talker's vowel space are the ones correctly perceived by the majority of listeners. It is important to note that the 'acceptable' areas are quite different from talker to talker. That is, it is not the case that listeners have entirely fixed notions of F1 and F2 values for each vowel, even when talkers have similar fundamental frequencies. The trend is for listeners to become more strict regarding steady state formant values as other articulatory information (e.g. transitional cues) is removed. That is, the 'correct' vowel spaces tend to shrink to some extent from the Sentence, to Word, to Gated Vowel conditions.

Table 4.25

Correlations Between Predicted and Obtained

Vowel Classification Scores

PREDICTIONS	CORRELATIONS WITH:	
BASED ON:	VOWEL CONDITION	WORD CONDITION
F1,F2	.7368	.7751
F1,F2,f0	.6988	.7371
F1,F2,f0,duration	--	.7185
'Best fit'	--	.7252

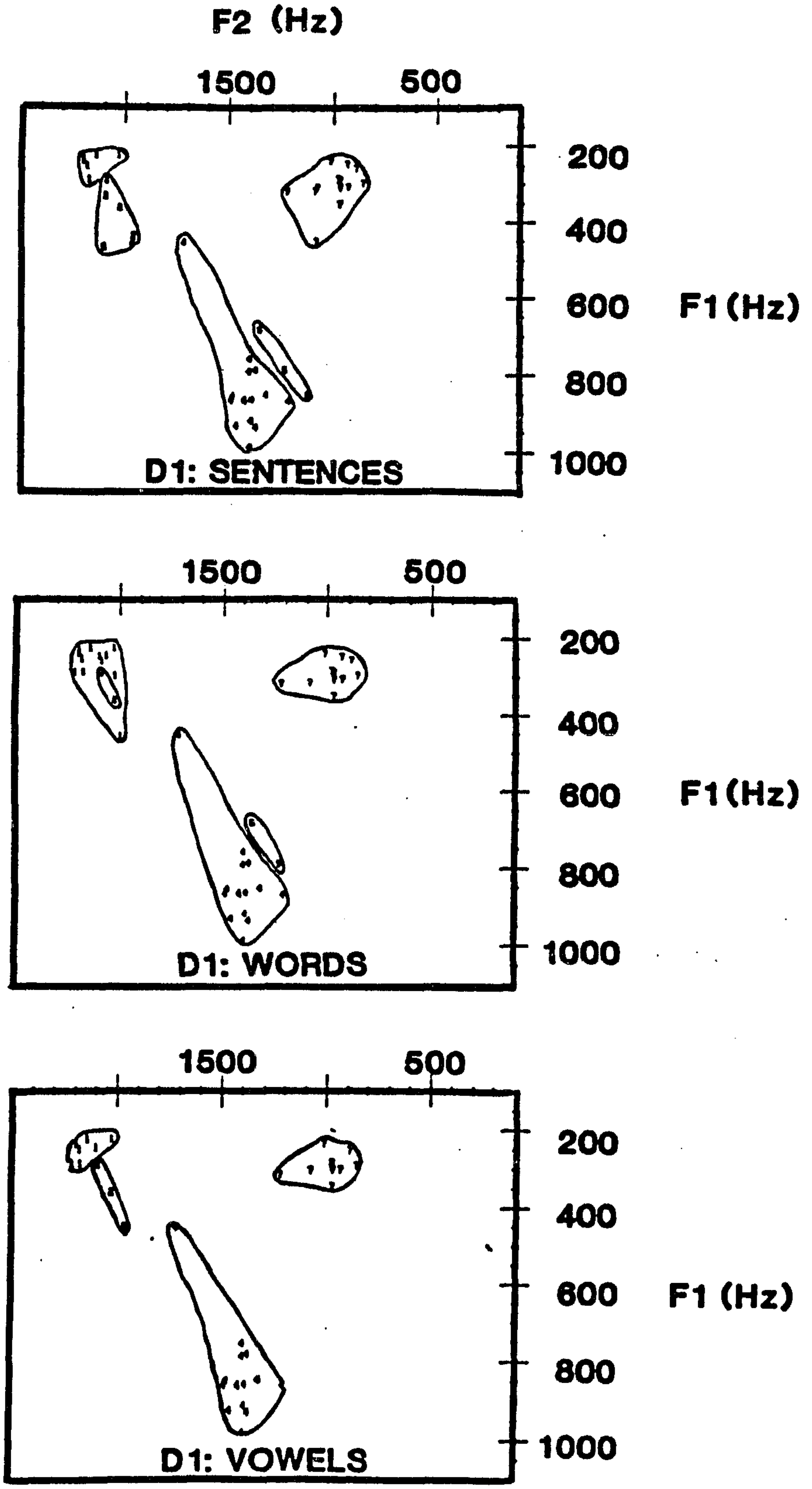

Figure 4.29 Token Plots for Correctly Perceived Vowels as a Function of Condition: D1.

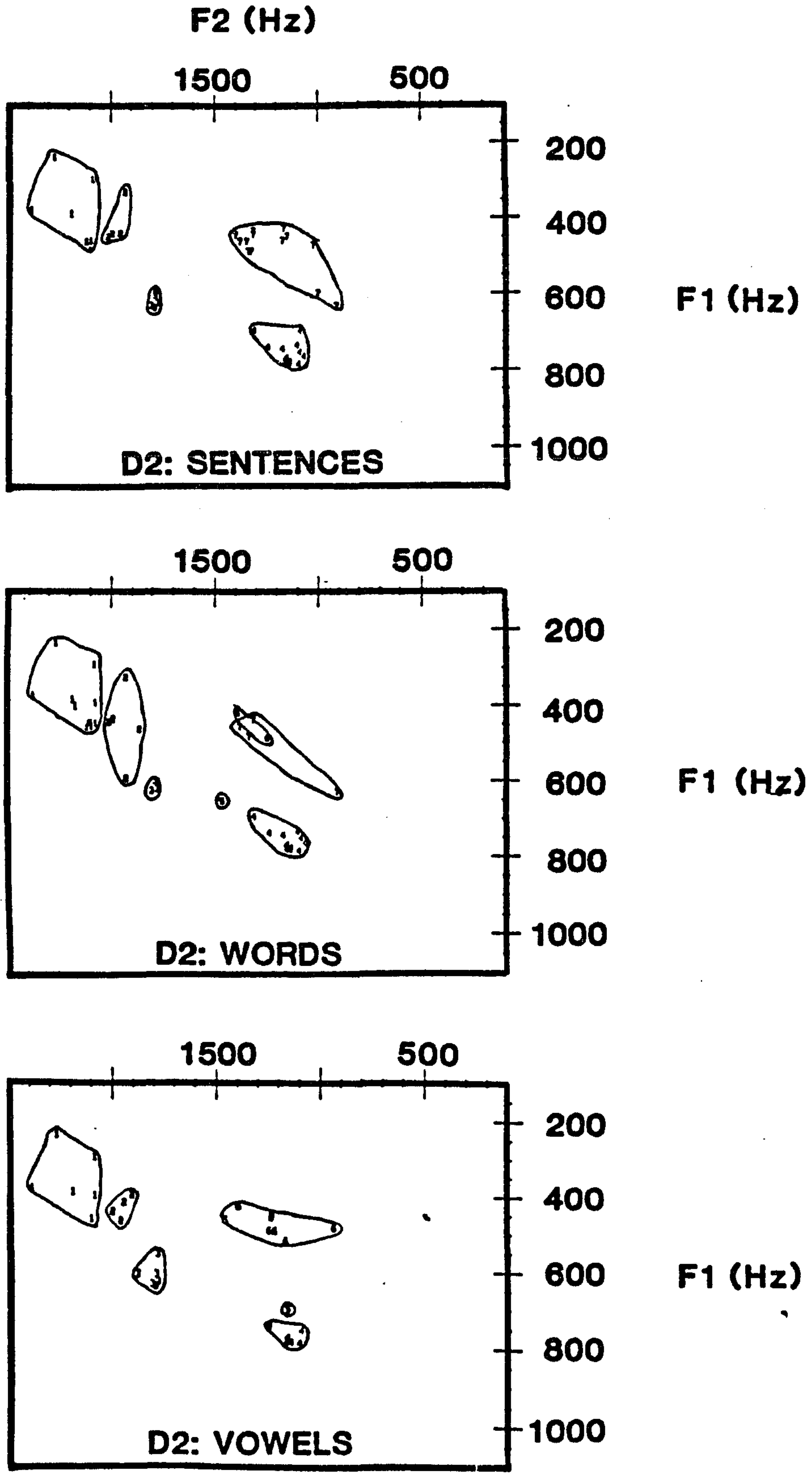

Figure 4.30 Token Plots for Correctly Perceived Vowels as a Function of Condition: D2.

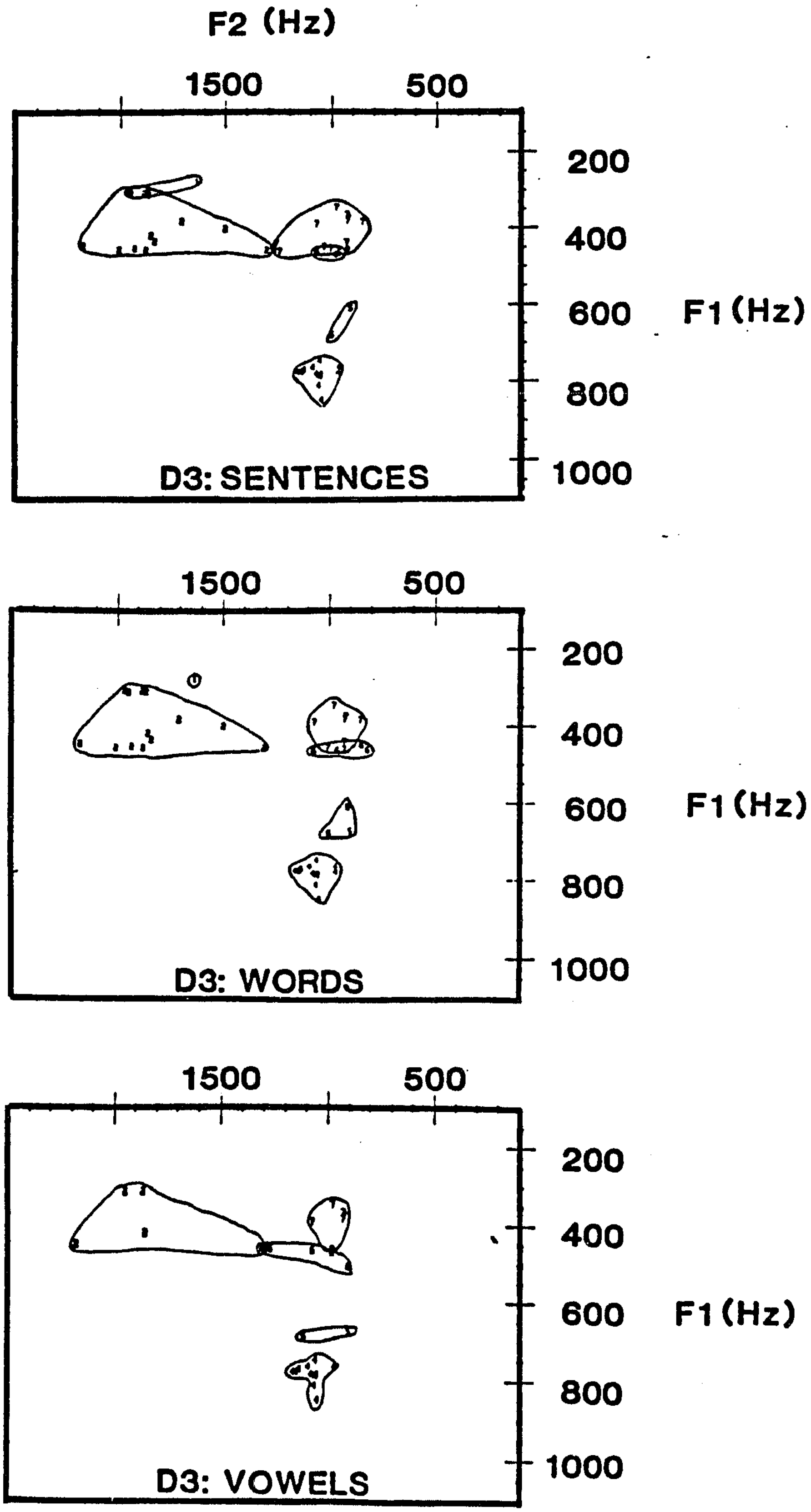

Figure 4.31 Token Plots for Correctly Perceived Vowels as a Function of Condition: D3.

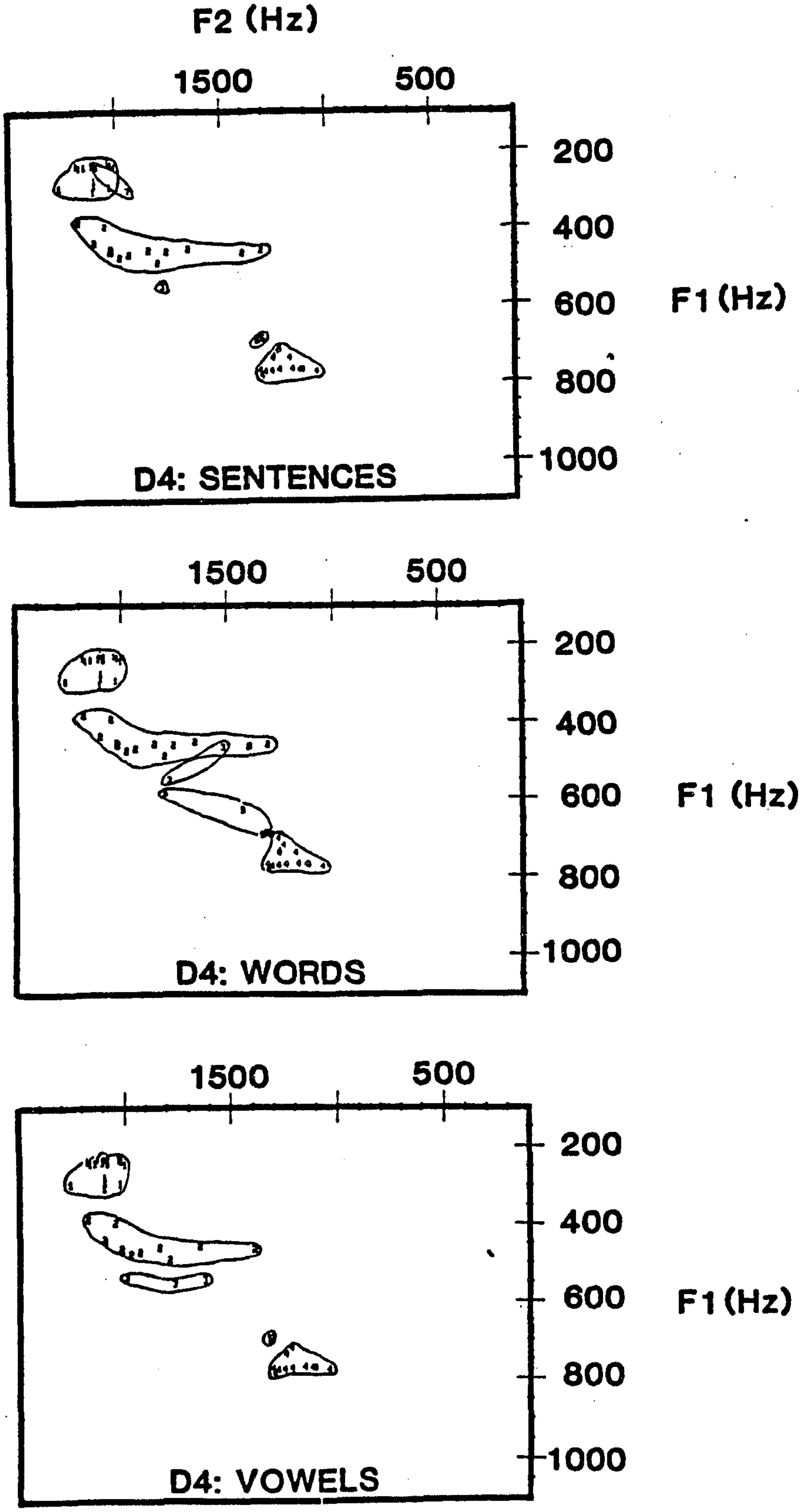

Figure 4.32 Token Plots for Correctly Perceived Vowels as a Function of Condition: D4.

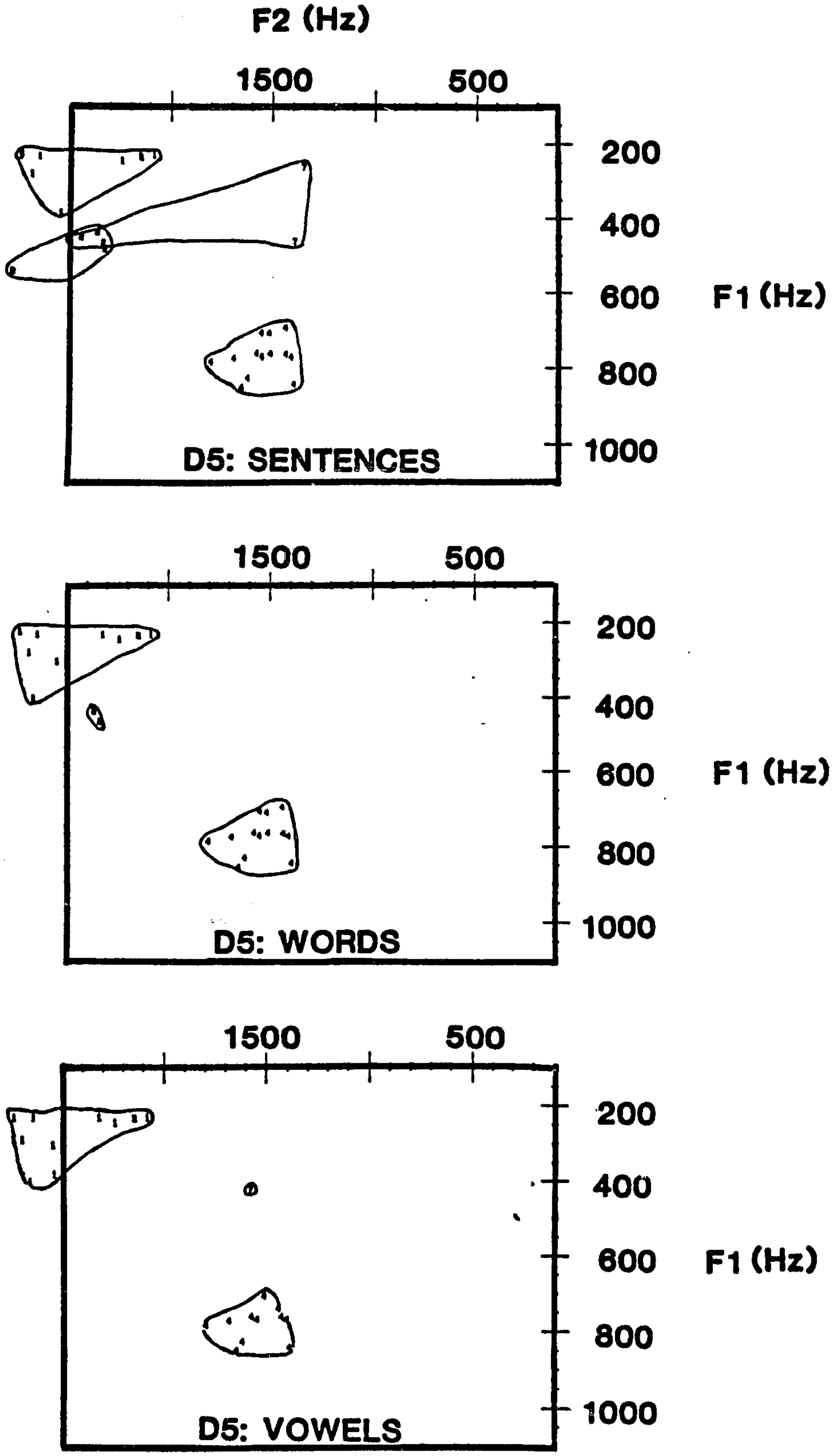

Figure 4.33 Token Plots for Correctly Perceived Vowels as a Function of Condition: D5.

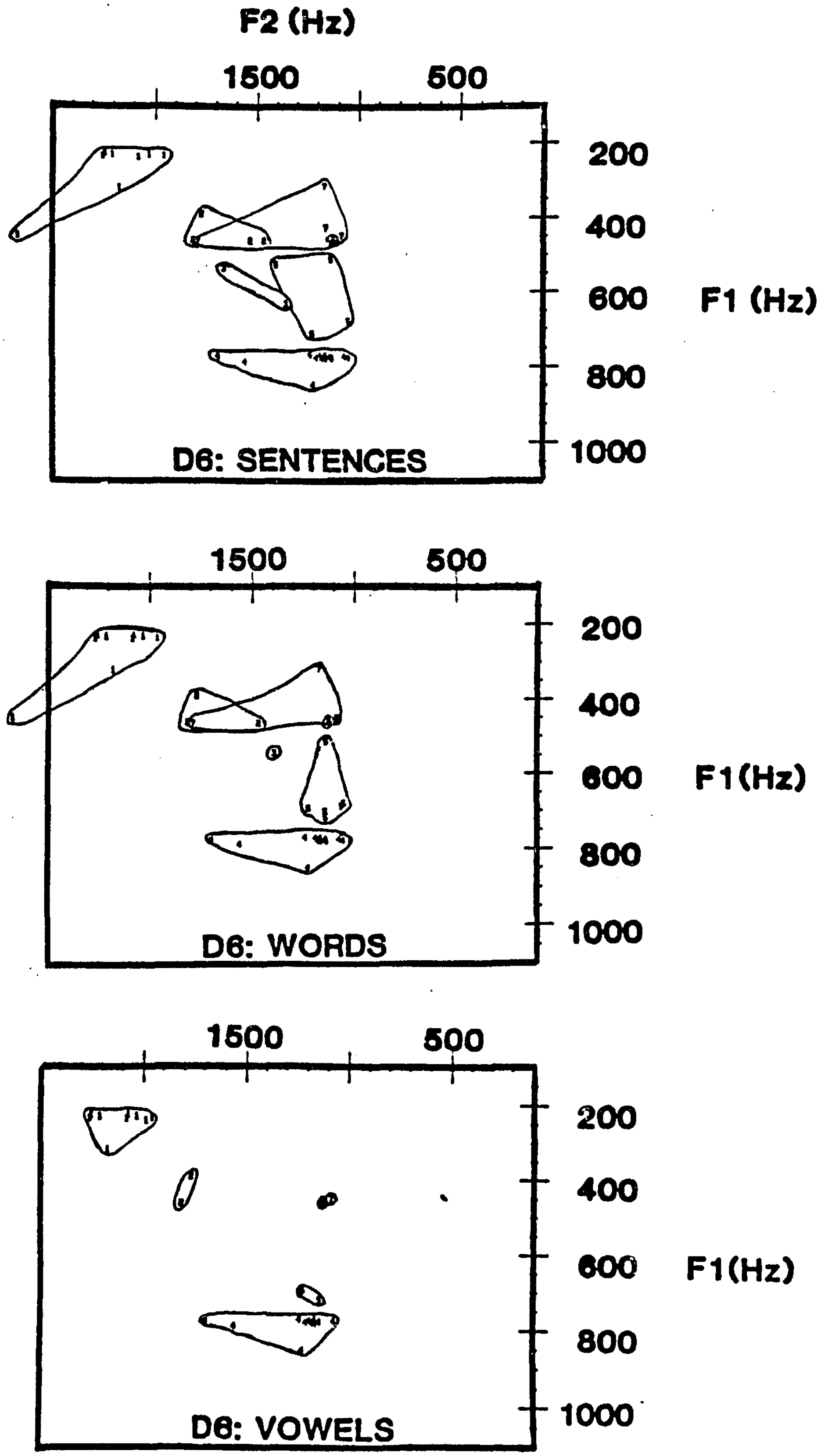

Figure 4.34 Token Plots for Correctly Perceived Vowels as a Function of Condition: D6.

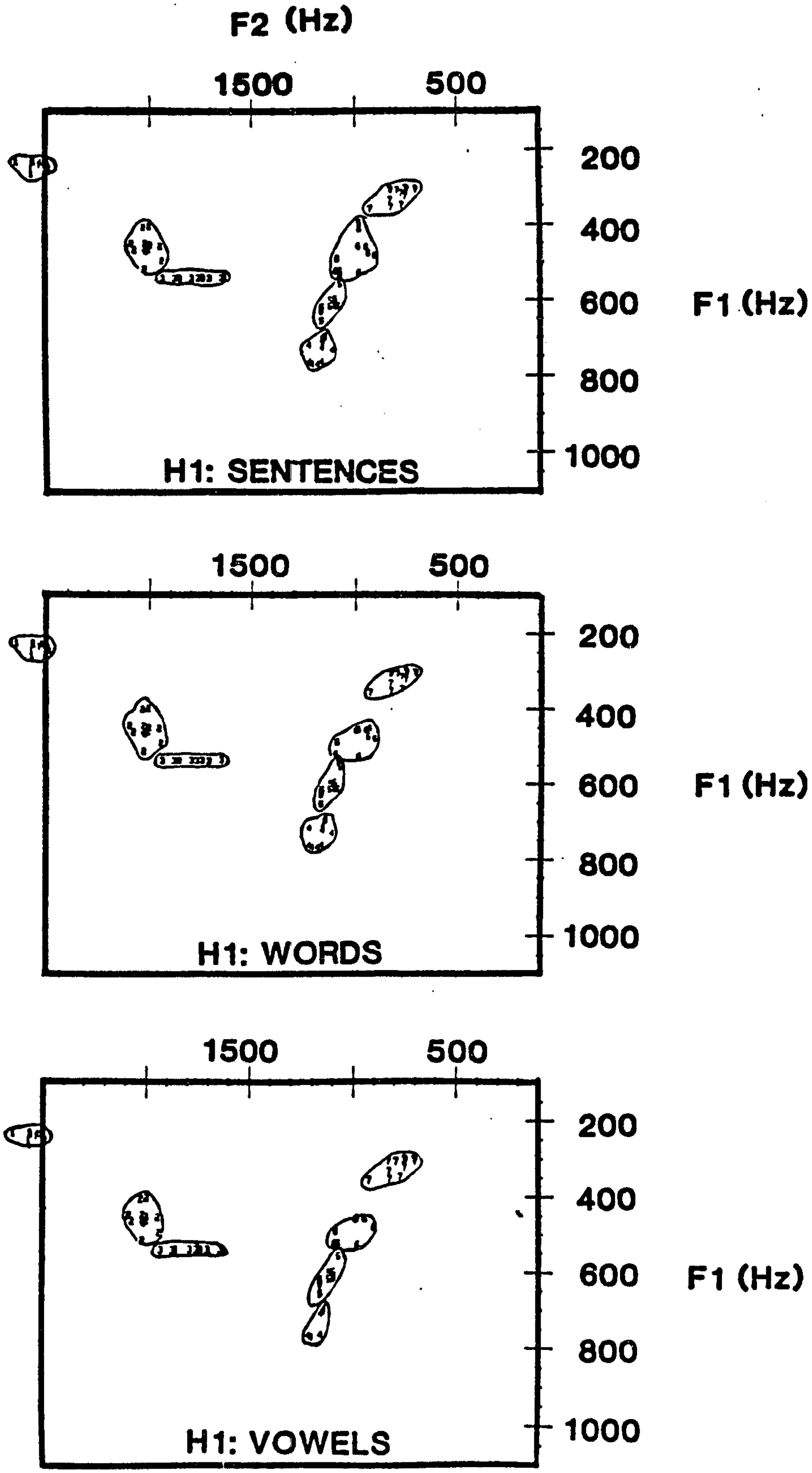

Figure 4.35 Token Plots for Correctly Perceived Vowels as a Function of Condition: H1.

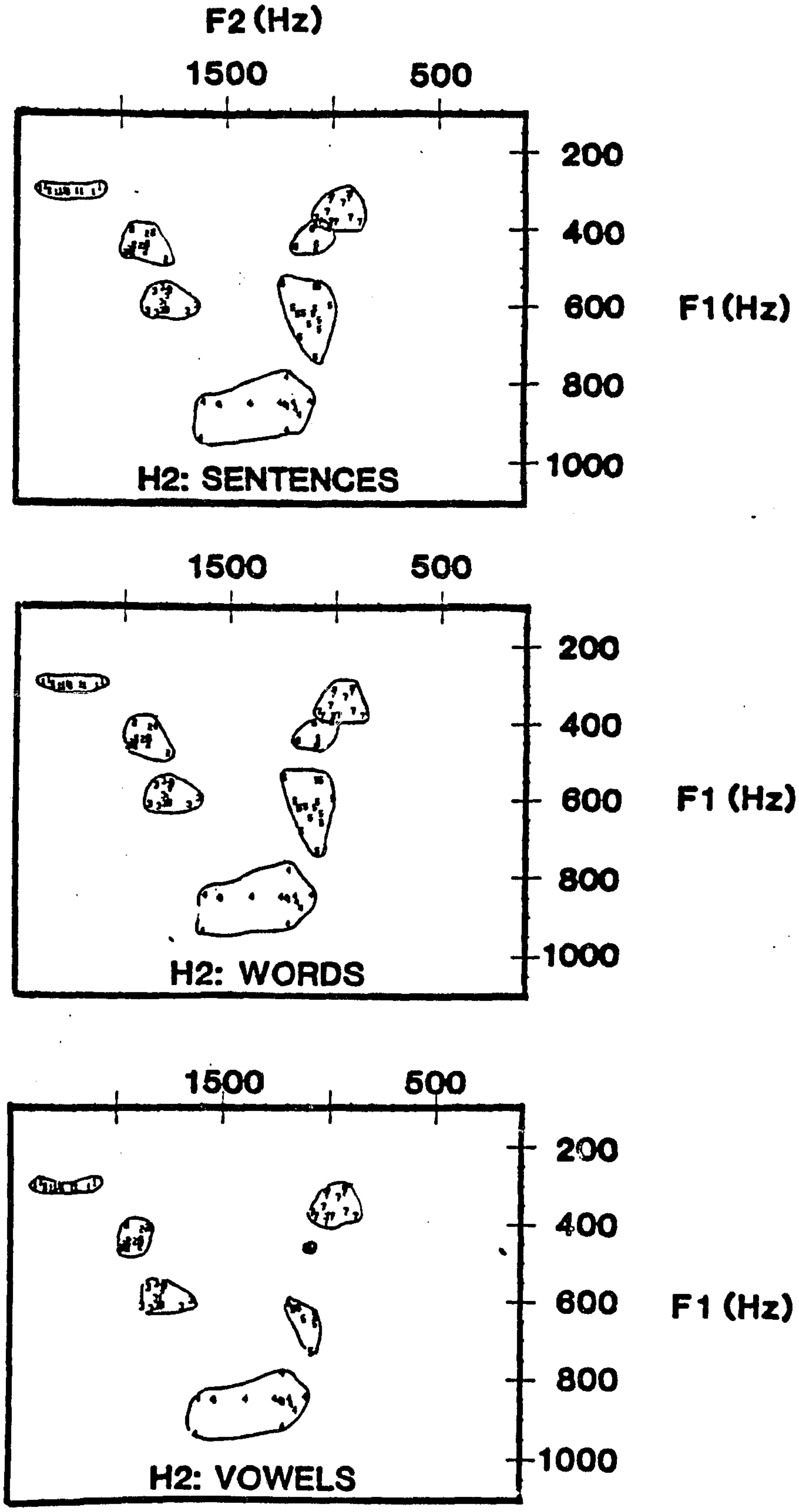

Figure 4.36 Token Plots for Correctly Perceived Vowels as a Function of Condition: D1.

FOOTNOTES

[9] Mean standard deviations for groups of talkers were determined by taking the average of each individual's standard deviation for the specified measure.

[10] Normalized data was determined as in Bush (1981).

[11] The perceptual significance of these results must be viewed with caution. Since the confusion matrices were used primarily for 'within-condition' analyses, they included those tokens that were removed for cross-condition analysis (see Chapter 3). Therefore, a portion of the error patterns in the vowel condition might reflect responses to stimuli which may or may not have been 'transition free'.

CHAPTER V

DISCUSSION & IMPLICATIONS FOR FUTURE RESEARCH

The principle goals of this study were to investigate intrasubject variability in vowel production by the hearing impaired, to test three hypotheses regarding the character of articulatory/acoustic vowel targets in this population, and to explore systematically the information used by listeners in perceiving vowels produced by the deaf.

ACOUSTIC ANALYSIS

Intersubject Variability:

The results of the acoustic analysis for grouped data (hearing vs. hearing-impaired talkers) are similar to those reported in the literature. Table 5.1 compares the group findings of the present study to similar measures reported by Angelocci et al. (1964) and Bush (1981).

Table 5.1

Ranges of of F1, F2, f0 and Duration as a Function of Vowel Target:

Comparison of Published Data

	NORMAL HEARING			HEARING IMPAIRED		
	A,K,& H (1964)	BUSH (1981)	PRESENT STUDY	A,K,& H (1964)	BUSH (1981)	PRESENT STUDY
f0	17 Hz	15 Hz	12 Hz	40 Hz	48 Hz	19 Hz
DUR	--	--	98 ms	--	--	124 ms
F1	655	388	523	330	405	462
F2	1715	1373	1517	1148	663	855

The trends for each of the acoustic ranges presented are qualitatively similar across studies. However, for three out of four of the acoustic measures in the present study (f0, duration, and F1), comparisons between the hearing and hearing-impaired group means do not adequately reflect the patterns shown by individual hearing-impaired talkers. For example, while the f0 range for the hearing-impaired group was 19 Hz (only slightly greater than normal), individual ranges varied from 4.2 to 35.7 Hz (that is, considerably greater than and less than that seen in the hearing controls). Moreover, within any given talker, one could not have predicted the range of one variable from the range of another. The danger inherent in arbitrarily averaging across deaf talkers has been noted by Bush (1981) and is supported by the present results. Interestingly, it would have been inappropriate to group talkers on the basis of the degree of hearing loss. While talkers D1 and D6 had the most profound losses of the group (107 and 105 dBHL pure tone average, respectively), they were quite different with respect to their patterns of vowel production (D1 was a Front/Back talker and D6 was an Overlapped talker). Talkers D3 and D5 also had similar pure tone averages (95 dB and 93 dBHL, respectively). As discussed in the previous chapter, D3 was one of the most proficient talkers with respect to vowel production while D5 was one of the least skilled.

Intersubject variability with respect to hearing vs. hearing-impaired talkers was such that we could not have made *a priori* assumptions based on normative data as to which vowel targets would represent the extremes of a talker's vowel space. For 5 out of 6 of the deaf talkers, as well as for the group as a whole, the lowest second formant frequency produced would have been overlooked had F2 range been arbitrarily defined as (F2 /i/ - F2 /u/). Had we based our analyses on these two vowels, the resultant ranges would have misrepresented the talker's vowel space. Therefore, in light of the intersubject variability in this population, it behooves researchers to analyze these talkers individually or, alternatively, to develop more appropriate methods of grouping hearing-impaired subjects. As will be discussed below, the results of the present study suggest that a measure of intrasubject variability in production deserves more attention as a possible basis for grouping hearing- impaired talkers (see discussion of Implications of Vowel Systems).

Intrasubject Variability and the Absent Target Hypothesis:

A number of studies have suggested that hearing-impaired talkers are highly variable and imprecise in their production of vowels (Angelocci et al., 1964; Monsen, 1976a; Smith, 1972). The apparent lack of control in production serves as the basis for the Absent Target Hypothesis regarding vowel production in this population.

While there may be excessive variability in vowel production by the hearing impaired, there has been no direct comparison between normal hearing and hearing-impaired talkers regarding this aspect of production. Table 5.2 shows the standard deviations for F1 and F2 in the four subject groups (H, A, B, C) as well as for individual talkers. Also shown is the published data for hearing talkers. It should be noted that the controls

Table 5.2

Variability in F1 and F2 Frequency:

Comparison to Published Data

	F1 Standard Deviation (Hz)	F2 Standard Deviation (Hz)
Potter & Steinberg (1950)	20-40	40-70
Eguchi & Hirsh (1969)	25	35-47
Pisoni (1978)	26	61
present study		
D1	89	113
D2	82	152
D3	65	157
D4	49	170
D5	95	281
D6	99	241
H1	25	53
H2	31	79
H	28	66
A	57	164
B	86	132
C	97	261

in this study agree quite well with the findings of other normative studies. The mean standard deviations (averaged across the 7 test vowels) of both F1 and F2 are greater for all hearing-impaired groups than for the normal controls. In each group, these mean differences prove to be statistically significant. This substantiates the notion that hearing-impaired talkers are more imprecise in vowel production than their hearing peers.

However, the Absent Target Hypothesis predicts a near random association between vowel type and formant frequency. The results of the Linear Discriminant analysis (see Table 4.11), based on F1 and F2 frequency for individual talkers, do not support the Absent Target Hypothesis. While the number of distinct targets is reduced to four or five for the hearing-impaired talkers, (as compared to seven in the controls) formant frequencies are not statistically random and articulatory targets cannot be considered absent.

Substitution of Control and a Deviant Phonology:

Physical and perceptual measures of vowel production might indicate a reduction in the number of distinct vowel targets because we are measuring or attending to the 'wrong' aspects of the signal. That is, when distinguishing between vowel targets, deaf talkers may substitute alternative parameters in production for those that are not perceptually salient and/or are difficult to control. Common examples given with reference to vowel production are the substitution of laryngeal adjustments for upper articulator movements (Angelocci et al., 1964; Horwich, 1977; Nickerson, 1975), and the exclusive use of duration to differentiate between the tense/lax pairs i/I and u/ʊ (Angelocci et al., 1964; McGarr & Gelfer, 1983; Monsen, 1974). The extreme form of this substitution process leads to Monsen's notion regarding a deviant phonology. This hypothesis holds that some hearing-impaired talkers abstract and operate within unique phonological systems for vowel production. Vowel differentiation, then, may be based on alternative articulatory/acoustic parameters, but production is well-controlled and targets clearly defined nonetheless.

As stated in Chapter II, this hypothesis implies near normal token-to-token variability in the non-conventional parameter(s). Additionally, it requires that these parameter(s) consistently distinguish between vowel targets in a way that the standard ones (F1 and F2 frequency) do not.

In this study, the alternatives considered were duration and fundamental frequency. While not the only possible substitutes for upper articulator movement, these two variables are the most frequently suggested candidates. The results of paired comparisons between groups of talkers (H,A,B,C) show all hearing-impaired groups to be significantly more variable than the normal controls for both of these measures. Therefore, consideration of the seven test vowels as a holistic system provides no support for the first criterion of a deviant phonology.

A weaker version of the Deviant Phonology Hypothesis might suggest that the hearing impaired use alternative parameters in specifying some but not all vowel contrasts. In fact, the results of the present analysis suggest that for certain vowels, variability in f0 and/or duration may be similar in hearing and deaf talkers. Linear Discriminant analysis was used to determine the differentiating power of these aspects of the signal alone and in combination with F1 and F2 frequency.

As expected based on the measures of overall variability, little if any increase in vowel differentiation is seen as the result of including alternative parameters in the discriminant function (see Table 4.11). Durational information differentiates between certain vowel pairs (i/I, u/ʊ, ɑ/ʌ) for three talkers where formant frequency did not. As has been suggested then, it appears that some hearing-impaired talkers substitute duration for fine articulatory adjustments in differentiating tense/lax pairs. As defined by Monsen (1976d) however, the use of a deviant phonology implies that the "speech of a particular deaf talker is not merely a collection of phoneme substitutions...". Instead, it is the "determined product of a phonological system" and is by no means "phonetically or phonologically inconsistent in itself" (p. 39). With this definition in mind, we should note that the three talkers who show the durational effect range from the most highly skilled (D3) to the least proficient (D6) in terms of formant related vowel differentiation. Therefore, the use of this alternative parameter is apparently unrelated to other articulatory skills. Furthermore, no talker shows this effect for all three contrasts mentioned. Thus, a pattern of system-wide substitution of control cannot be found in any individual. One would have to question whether the results are indicative of deviant phonological systems as just defined. Instead, the durational contrasts noted here may be better described as isolated production phenomena.

The results of talker D6 warrant special mention. This subject alone demonstrates the necessary components of a deviant phonology. He exhibits high variability in F1 and F2 frequency, low variability in an alternative parameter, and an improvement in vowel resolution as the result of including f0 and duration in the discriminant function. However, the three vowel contrasts added all appear to depend on a different variables or combination of variables. /I/ vs. /ɛ/ is specified by f0, /ɑ/ vs. /ʌ/ is specified by duration and /ʊ/ vs. /u/ is specified by both. /i/ vs. /I/, unexpectedly, is specified by formant frequency alone. The lack of consistency within this vowel system makes it difficult to argue that this speaker has abstracted an internally consistent deviant phonology for vowel production.

The notion of a deviant phonology in vowel production could have been substantiated by certain patterns in the perception of these utterances. Specifically, one would have predicted that a priori knowledge of the patterns of deviation in deaf vowel production would have afforded the listener a significant perceptual advantage. The absence of an experienced listener effect, however, argues in favor of the notion that hearing-impaired talkers operate within systems which can be defined by the normal variables, and are, as such, available to the naive listener. Instead, one might suggest that the characteristic features of a deviant phonology are specific to individual talkers and not to deaf vowel production in general. If this were the case, we would predict a significant effect of being given the opportunity to sample the individual's deviant phonology directly (that is, a significant calibration effect). However, the calibration effect was not significant in the perception of deaf vowels. Finally, one could argue that the experienced listener would adjust perceptual criteria only if sufficiently oriented to the task (that is, realize that the speaker was deaf). However the relationship between correctly identifying the speaker as hearing impaired and correctly identifying the vowel was not significant for either group of listeners.

Thus, we have both acoustic and perceptual evidence which strongly support the notion that the hearing impaired operate within the normal phonological system, at least with respect to vowel production.

Vowel Systems in the Hearing Impaired:

The results of this study do not support the Absent Target or the Deviant Phonology Hypothesis of vowel production in the hearing impaired. The following qualifications should be noted. With respect to the Absent Target Hypothesis, one should remember that the talkers in this study did not include those judged to be completely unintelligible. There may be evidence of an absence of vowel targets in this subset of deaf talkers. With respect to the Deviant Phonology Hypothesis, it may be the case that by using a different group of talkers (eg. those demonstrating extreme hypernasality) and/or a different set of articulatory variables (eg. relative nasality, degree of breathiness) one could find support for the Deviant Phonology Hypothesis. These issues should be explored fully before rejecting either of the above hypotheses.

Nevertheless, the results of the present study support the 'hybrid' hypothesis discussed briefly in Chapter I. That is, the findings of this study suggest that hearing-impaired talkers function within vowel systems which differ from the norm in the number of discrete targets and to some extent in target locations. They do not differ in the target variables themselves, however, distinguishing what we might call a 'Reduced Target Hypothesis' from the Deviant Phonology Hypothesis. The question is how best to describe the reduced vowel systems observed.

In the case of hearing talkers, mean values of F1 and F2 as a function of vowel target can be used to describe vowel systems fairly completely. Individual targets vary to some extent regarding token-to-token variability and information about standard deviations of the first two formants is of theoretical interest. These data would not, however, substantially alter the characterization of a given talker's system for vowel production. For example, the token plots for the controls (Figs. 4.15 and 4.16) are not particularly revealing.

The same is not true for the vowel systems of the hearing impaired. In fact, the token vowel plots shown in Figs. 4.9-4.14 are considerably more informative descriptions of the vowel space than are the mean plots in Figs. 4.1-4.6. Two main sources of information are gained. First, token plots allow for a comparison between vowel targets in terms of articulatory control in production. Mean plots do not. While this may be a trivial observation with respect to mean vs. token measurements in any data set, this comparative feature of analysis turned out to be crucial to the characterization of the vowel systems observed. Secondly, token plots, or more precisely measures of intrasubject variability, are essential if estimates of formant range are to be informative. That is, a large formant range may be indicative of appropriate articulatory movement in vowel differentiation or articulatory movement which is uncontrolled and excessive. Alternatively, narrow formant ranges may result from immobile articulators and neutralized vowels or from a well-defined but restricted vowel space (as in Talker D2 in Stein, 1980). Data concerning intrasubject variability clarify this kind of ambiguity. Three styles of vowel production were delineated from token plots:

1. Point Vowel Systems
2. Front/Back Systems
3. Overlapped Systems

Articulatory interpretations of acoustic data must be considered highly speculative. However, such theorizing provides interesting directions for future research using physiological measures.

For the Point Vowel System (as exhibited by talkers D3 and D4; Figs. 4.11 and 4.12), two to three tongue positions may be hypothesized: 1. high front 2. high back and/or 3. neutral. The last of these may have served for the production of /ɑ/ in combination with lowering the jaw. For talker D4, acoustic measurements at the center of the vowel /u/ indicated the production of a high second formant (as shown in Fig. 4.11), as did the predominance of /i/ responses in the perception of this vowel type. However, most of these tokens would have been narrowly transcribed as a high front rounded vowel or the diphthong /iu/. Thus, it might be that this talker never used a tongue-backed position and simply employed liprounding to differentiate /i/ from /u/. For both D3 and D4, tongue position for the non-point vowels appeared to be undefined in the front/back dimension although these vowels may have been differentiated to some extent in terms of tongue height. Alternatively, F1 movement may have been controlled by jaw opening alone. The excessive use of jaw movement for vowel differentiation has been suggested by others (Ling, 1972; Martony, 1968) and verified by Stein (1980) with physical measurements. More work needs to be done to establish the exact nature of this production strategy.

For the Front/Back System (as exhibited by talkers D1 and D2; Figs. 4.9 and 4.10), tongue height was clearly an irrelevant dimension. High vs. low first formant frequencies were probably controlled by gross changes in jaw height alone. However, these talkers unquestionably demonstrated their ability to make a constriction in the front vs. the back of the oral cavity. This is somewhat unusual in the deaf population since, as we know, tongue movements are to a large extent invisible and inaudible to most deaf talkers. Although D1 and D2 were not the most intelligible talkers in the group, their mastery of this relatively difficult articulatory skill may be indicative of their potential for highly intelligible speech.

In the Overlapped System (as exhibited by talkers D5 and D6; Figs. 4.13 and 4.14), articulatory control was obviously poor. A significant feature in this style of production was the relatively large range of formant values seen. That is, one could not conclude that these talkers failed to differentiate between vowels as the result of minimal articulatory movement. On the contrary, their vowel production is to be considered inadequate because they appear to employ uncontrolled articulatory gestures. This is an unexpected finding with respect to the commonly accepted notions of vowel neutralization and vowel reduction in the hearing-impaired population. Concurrent acoustic and physiological measures would be appropriate directions for future research. Such experiments might help uncover the physiological 'constant' in a gesture whose acoustic consequences are highly variable.

If these three descriptive systems could be used to categorize deaf vowel production in general, it would be unnecessary to measure a full array of vowel targets for the purposes of classification. One of the high point vowels (/i/ or /u/) and two to three of the intermediate vowels (e.g. /I,ʊ/ or /ʊ,ɛ,I/) could completely specify the system by delineating (1) the point vs. intermediate vowel contrast and (2) variability in the front/back dimension. However, in light of the excessive intersubject variability in this population, one must be cautious in generalizing the results of this study to the population as a whole. A large number of hearing-impaired talkers would need to be evaluated on the full array of vowels in order to consider these classification schemes complete.

Nevertheless, results strongly suggest that for some talkers, the production of point vowels is less variable than the production of intermediate vowels. The reduced acoustic variability noted here for point vowels may be explained, in part, in terms of the 'quantal theory of speech production' (Stevens, 1972) rather than in terms of true differences in articulatory control. That is, as the result of non-linearities in articulatory/acoustic relations, equal amounts of articulatory variability at different constriction locations (e.g. point vs. non-point) may result in unequal amounts of acoustic variability in the output signal. However, if this were the sole explanation for the effect seen here, one would expect to see the same degree of differentiation between tense and lax vowel variability for all of the talkers. As has been noted, this was not the case. For some talkers then, there appear to be real differences in the articulatory control for point vs. intermediate vowels. In understanding why this may be so, one should remember that the extremity of the articulatory gestures which would be considered appropriate for point vowel production may make these targets easier to teach and master in terms of visual, tactile and kinesthetic feedback.

Implications of Vowel Systems and Variability in the Speech of the Hearing Impaired:

It is clear from the data presented that these hearing-impaired talkers move their articulators for vowel differentiation. While some of the vowel contrasts appear to be neutralized, the vowel space is not. Point vs. non-point distinctions may guide one system while the front vs. back distinction governs another. In all cases, variability is greater than normal and this argues against the notion that the deaf simply do not move their tongues for vowel differentiation.

A number of investigators have reported inappropriately high second formants in the production of back vowels (Angelocci et al., 1964, McGarr & Gelfer, 1983; Stein, 1980; Suonpaa & Aaltonen, 1982) as well as physiological evidence of tongue fronting (McGarr & Gelfer, 1983; Stein, 1980). Acoustic measurements in this study indicate that the second formant frequency was rarely as low as that seen in normally produced back vowels (see Table 4.4) and thus support the notion of a tongue-fronted posture. For some talkers, liprounding may have been the only factor in F2 lowering for back vowel targets. As Nober (1967) suggested, the relatively high intelligibility of back vowels noted for deaf talkers may be more a result of appropriate liprounding gestures than correct tongue postion.

Intrasubject variability proved to be one of the most obvious and descriptive characteristics of the vowel systems observed. Some of the implications of this aspect of production have already been discussed. In addition to its theoretical significance, token-to-token variability may prove to be a useful grouping variable. That is, the groups formed on the basis of intrasubject variability in vowel production (A,B,C) also serve to group the talkers in terms of vowel intelligibility (see Table 4.13) and in terms of overall speech intelligibility (see Table 3.1). Formant ranges, however,could not have been predicted. Large ranges co-occured with both the most and least variable talkers. A measure of variability alone then, suffers from some of the same ambiguities as does the measurement of formant range. A numerical index which combines these two sources of information would be unique with regards to its descriptive power. The following equation yields a 'vowel skill index' (VSI) which may be a useful grouping variable for future research:

$$\text{F1 range/F1 st. deviation} + \text{F2 range/F2 st. deviation} = \text{VSI}$$

Table 5.3 shows the ranges, standard deviations, and VSI's for each talker and for the four subject groups. The VSI's cleary separate the controls from the hearing-impaired groups, and Groups A and B from

Table 5.3

F1 and F2 Ranges, Standard Deviations, VSI's and Flag Values for Individual Talkers

Talkers	F1 Range (Hz)	F1 Standard Deviation	F2 Range (Hz)	F2 Standard Deviation	VSI	Flag Value
D1	500	89	880	113	13.41	32
D2	336	82	934	152	10.24	38
D3	498	65	909	157	13.45+	104
D4	566	49	1147	170	18.30+	170
D5	431	95	859	281	7.60	18
D6	440	99	994	241	8.56	40
H1	561	25	1277	53	46.53	16
H2	486	31	1758	79	37.93	46
H	524	28	1517	66	41.71	35
A	418	57	1758	164	18.05+	137
B	532	86	1028	132	13.98	29
C	436	97	927	261	8.04	31

Group C. As it stands, however, the index does not discriminate those speakers with large differences in standard deviations for different vowels from those without. The results of the following equation can serve as a qualifier for the VSI's:

$$\text{F2 st. dev.(non-point vowels)} - \text{F2 st.dev.(point vowels)} = \text{Flag}$$

If the flag value is greater than ±50 Hz, we will mark the VSI with a '+' sign; the VSI will remain unmarked if the flag value is less than ±50 Hz. By so doing, the Point Vowel talkers become distinguishable from the Front/Back talkers by virtue of their positive flag. Both groups are relatively highly skilled in comparison to the Overlapped talkers (as determined by their high VSI's) . The flag, however, would indicate that there are qualitative differences in the production styles of these two groups of talkers.

Large token-to-token variability has now been noted in a variety of studies and for a number of different aspects of production. Researchers should question the validity of measurements on the hearing impaired which are based on a small number of repetitions. Similarly, therapists and teachers who evaluate the speech of the hearing impaired would do well to incorporate repeated measures in diagnostic evaluations. Even if variability is only noted perceptually, this aspect of production may prove to be useful in planning teaching strategies. The child who appears to 'have' a phoneme, then 'lose' it, then 'have' it again, may simply be demonstrating his/her inability to control articulatory gestures and not the acquisition and loss of phonemic entities. Moreover, the common problems associated with 'carryover' outside the classroom cannot be adequately addressed until we understand the nature of the variability in production seen in the teaching situation. It may be that we should use a different instructional approach for talkers who can demonstrate an ability to control articulatory movements than for talkers who cannot. Also, knowledge of the current system being used by the child should allow the teacher/therapist to focus work on the most deficient contrasts.

PERCEPTION OF DEAF VOWELS

Vowel perception was evaluated under conditions providing controlled amounts of dynamic information. The question was simply: Do experienced and/or inexperienced listeners use the same sources of acoustic information in disambiguating vowels which have been spoken by hearing-impaired as compared to normal hearing speakers.

As mentioned earlier, a specific interest was whether *a priori* knowledge of the patterns of deviation in deaf speech allow listeners to 1. change strategies (e.g. ignore certain sources of information or attend to novel ones) or 2. make better use of diminished cues or 3. increase overall performance. The perceptual experiment provided a forum for evaluating these details of the listener effect.

Comparison to Published Data:

As discussed in Chapter II, methodological differences between studies make it difficult to compare reported error rates directly. However, the relative intelligibility of vowels can be discussed. In this study, the rank ordering of vowels re. intelligibility followed the pattern suggested by Suonpaa & Aaltonen (1981). That is, the point vowels generally showed the lowest error rates. It was not the case, as has been suggested before, that back vowels (Gold, 1978; Mangan, 1961; Nober, 1967; Smith; 1972) or low vowels (Nober, 1967; Smith, 1972) were, as a group, more intelligible than front or high vowels.

As suggested by Smith (1972) there is an obvious and direct relationship between overall and vowel intelligibility. One point must be made in this regard. Given his poor overall intelligibility, talker D6 appears to be unexpectedly high in the ranking of vowel intelligibility. This talker produced the largest number of 'short' tokens which had to be removed from analyses involving condition effects. With these tokens in the analysis, his scores drop precipitously in all three conditions (leaving the difference scores equivalent). For this talker then, it appears that the short tokens were poorly produced as well. This pattern was not found for any of the other talkers' unadjusted scores.

Sources of Information for Vowel Perception:

The results of this study show that the calibration effect is not significant in the perception of vowels produced by hearing or hearing-impaired speakers.

With respect to the normal literature, this does not support Joos's (1948) original claim that listeners calibrate to the talker's vowel space. The findings with regards to deaf vowels are of particular interest. These speakers differ from normal hearing talkers in both rate and vowel space. Thus, deaf speech provides an ideal forum for a test of the calibration hypothesis. The absence of such an effect fails to support the notion that listeners are able to judge vowel identity in relation to the tokens immediately preceding the test signal.

As discussed earlier, the absence of a calibration effect in the speech of the deaf also argues against the notion of a well-defined but deviant vowel space. The signficance of the Listener x Condition x Vowel interaction suggests that, in relative terms, naive listeners benefit from information in the precursor sentence in tense vowel recognition, whereas experienced listeners benefit slightly in lax vowel recognition. It may be that the two groups attempt to use the information in the carrier phrase differently. Naive listeners may attempt a direct mapping of the test vowel onto the exemplars in the carrier phrase. Experienced listeners may judge ambiguous test vowels in relation to the vowel space just constructed.

Listeners were afforded a significant perceptual advantage in the identification of hearing as well as deaf vowels when provided with transitional information. There was no statistically significant interaction between Group and Condition. This indicates that the informational value of the surrounding acoustic signal was unaffected by the hearing status and/or vowel system of the speaker.

With respect to the normal hearing talkers, this finding confirms the results of other Type A vowel perception studies (that is, studies which compare the intelligibility of vowels produced and heard in CVC syllables to portions of the test vowels gated out of their original context). The size of the effect is, however, somewhat smaller than previously reported for comparable conditions (Assmann et al., 1982; Ochiai & Fujimura, 1971). This may be seen largely as the result of an inflated error rate for the CVC syllables in the present study. One might ask why the CVC error rate was so high in this experiment. It may be the case that the normal hearing controls were less intelligible than the average adult talker. This is somewhat improbable however. More likely, the difficult nature of the test situation (mixed hearing and deaf speakers) decreased performance in what otherwise would have been an easy listening condition.

The statistically significant interaction between Vowel and Condition reported here is not unique. The majority of studies in the literature have not included 'vowel' as a factor in their analyses. However, those that report the data for individual and/or groups of vowels also show differences in the magnitude and/or direction of the effect (Assmann et al., 1982; Gottfried & Strange, 1980; Macchi, 1981; Strange et al., 1979, 1976).

With respect to the hearing-impaired talkers, the finding of a context effect speaks directly to theories concerning how listeners decode the speech of the deaf. Listeners clearly demonstrated the ability to extract relevant information from formant transitions. These resonances reflect articulatory gestures towards and away from a steady state target and as such represent dynamic aspects of articulation. It may be that the movements are guided by mistimed motor commands or are simply slow and laborious. However, if a source function is operative, it must be the case that the articulatory gestures generate acoustic consequences in the speech signal. The results of this study suggest that listeners use this source of information, as they do in normal vowel perception, to disambiguate vowel targets.

A potential difficulty in ascribing the significance of the context effect solely to the elimination of transitional information was the concomitant equating of vowel duration. That is, the gated vowel files were all 100 ms. in duration while the vowels in the Word condition varied freely along this dimension. If this factor were entirely responsible for the context effect, one would have to assume that listeners were effectively using durational information for vowel perception. This notion and the possibility that the context effect was solely the result of removing durational cues is not supported by a number of findings in the data:

1. Steady state formant values for deaf vowels were frequently ambiguous, so the high error rate in the Gated Vowel condition was not surprising. If hearing the vowel in its original consonant context added durational cues only, one would predict simply a decrease in near

neighbor (e.g. tense/lax) errors. In fact, we observe a 66% decrease in cross-triangle confusions. This suggests that listeners were benefiting from the inclusion of transitional information.

2. Vowels produced by the hearing-impaired talkers were longer than those of the controls. If listeners were relying on durational cues for vowel identification, it would surely have been helpful (if not necessary) to judge duration relative to the characteristic durations of a given talker. That is, we would expect a significant calibration effect. The absence of this effect argues against the notion that listeners were, in fact, using durational information for vowel identification.

3. On the other hand, what if cue value rested in absolute duration? In this case, we would not expect an improvement in scores to result from the addition of durational information to deaf vowels since the hearing and hearing-impaired version of a given vowel target often differed in absolute duration by more than any such contrast carrying phonetic value. Being exposed to such misleading cues in the Word Condition should not have increased performance. This, of course, was not the case.

4. Although it is highly improbable, one could suggest that listeners were in some way able to judge the absolute duration of deaf vowels against some internally generated deaf standard. Again, this would force one to predict a significant relation between correctly identifying the speaker as deaf and correctly identifying the vowel. As discussed earlier, this relation was not significant.

Therefore, a number of findings indicate that the consonant context effect was, at least in large part, the result of manipulating the amount of transitional information in the acoustic signal.

The finding of a context effect in both groups of talkers, the uniform absence of a calibration effect, and the absence of an experienced listener advantage all converge on the notion that listeners go about the task of vowel identification in a way that is unaffected by the hearing status of the talker. There is no evidence to suggest that listeners are prevented from decoding the deaf signal in the 'normal' way nor that they develop new strategies as the result of experience. The absence of a significant Listener x Condition interaction suggests that no qualitative differences in perceptual strategy result from experience with deaf speech. The two groups of listeners appear to 'look for information' in the same places and to use the information with the same effectiveness.

The fact that the magnitude of the context effect was equivalent for deaf and hearing talkers is somewhat surprising. A few qualifications seem warranted.

First, it has frequently been reported that bilabial plosives are one of the simplest consonants for hearing-impaired children to produce. It may be that the context effect would have been reduced had a different consonant environment been chosen.

Second, vowel intelligibility for the normal talkers in the Word condition was poorer than expected. The use of a blocked speaker design might improve the Word Condition scores for hearing talkers only. If this were the case, there would be a significant interaction between Group and Condition, suggesting a difference in the informational value of formant transitions.

Finally, phonetic contrasts calling for more temporal precision in their execution (e.g. voiced/voiceless; glide/stop; fricative/affricate) may prove to be more sensitive to the perceptual effects of poor coarticulatory skills in the hearing impaired than vowel contrasts. Since formant transitions carry both consonant and vowel information, a similar experimental design could be used to test whether transitions in the speech of the deaf specify consonant as well as vowel targets. All of the above represent avenues for future research.

In short, the results of the present study suggest that deaf talkers produce a signal in which transitional information accurately specifies articulatory movement, and that listeners are able to use this information effectively in decoding the signal (at least within the limited context of a /bVb/ syllable). However, more work needs to be done before we can reject the notion that poor coarticulatory skills play a role in the unintelligibility of deaf speech.

Recognizing the Deaf Voice:

The factors which lead to the recognition of a speaker as hearing impaired have been largely unexplored. This at first may appear to be of only cosmetic concern. It is, in fact, an important issue with respect to the underlying deviant qualities of the speech signal produced by these talkers. Calvert (1961) reported that listeners require a sample of speech long enough to include articulatory movement in order to identify the speaker as deaf. This ascribes a great deal of perceptual salience to the deviant nature of articulatory gestures and very little to the source function. A number of other investigators have shown that the deaf source function is quite different from that of normal hearing talkers and suggest that it is a significant factor in the abnormal quality and unintelligibility of their speech (Forner & Hixon, 1977; McGarr & Osberger, 1978; Stevens et al., 1983; Whitehead, 1983).

Results from the present study support Calvert's findings that the majority of information leading to the perception of the deaf voice is coded in articulatory information as a function of time. The effect of Condition on listeners' scores in this regard was highly significant. However, it was not the case that the Gated Vowel condition was devoid of information relevant to the task. That is, listeners performed significantly better than chance given steady state information only. Thus, some resonance and/or source function characteristics peculiar to the speech of the deaf were apparently present and perceptible in this condition. A number of acoustic variables could have been carriers of such information and warrant further investigation. The proximity and/or relative values of f0, F1 and F2 frequency, the bandwidth and amplitudes of formants, the periodic character of the source function, the presence of excessive nasality and the average fundamental frequency are the most obvious candidates. It would be reasonable to assume that the importance of average f0 was reduced in the present study by the exclusion of talkers with excessively high pitched voices. Nevertheless, the mean difference between the two groups in f0 was 7.1 Hz and individual talkers ranged from being 2 Hz to 22 Hz above the controls in this regard. Moreover, there is some evidence to suggest that above chance performance in the Gated Vowel condition cannot be explained solely in terms of perceived nasality. The evidence derives from the fact that the presence or absence of nasality in the signal (as determined from the 'nasality files') did not predict which vowel spectra would be technically difficult to analyze. That is, many of the tokens which could not be analyzed with LPC techniques showed no evidence of excessive nasality. Other source-related factors were apparently responsible for ambiguities in the vowel spectra and therefore may have led to the perception of the 'deaf voice'.

An interesting approach for future research would be to control, artificially, the aforementioned parameters in a synthetic signal whose details were otherwise controlled by the hearing-impaired speaker.

As in the phonetic task, there does not appear to be a significant difference between experienced and inexperienced listeners in identifying the hearing status of the talker. That is, despite years of experience with these talkers, teachers of the deaf are no better able to recognize a speaker as hearing impaired than are their naive peers. There was no interaction between Condition and Listener. That is, even when given steady state information alone, an untrained listener performs similarly to a highly trained peer. It should be remembered that the task required only that the listener identify normal vs. non-normal speech. Perhaps a difference would have been noted had the task required differentiation between deaf speech and other types of clinically disordered speech (e.g. aphasic speech).

LISTENER EXPERIENCE

A good deal of disagreement exists in the literature regarding the presence or absence of an experienced listener advantage in the perception of deaf speech. Two studies address the listener issue specifically with regards to vowel perception (Gulian & Hinds, 1981; McGarr & Gelfer, 1983). The results of the present study do not show a listener advantage and thus agree with the findings of Gulian & Hinds, but not with those of McGarr & Gelfer. One can conclude that experienced listeners are no better than their naive peers at calibrating a deaf vowel space, at least under the conditions of the present study. Nor are they able to extract additional information from formant transitions. The equality of the two groups is stable across talkers. That is, there is no Listener x Group interaction. Therefore, it cannot be said that the presence of a listener advantage is a function of the overall intelligibility and/or variability in vowel production. How can we explain the contradictory results of studies in this area of research?

First, an interesting trend in the literature is that those studies using a blocked speaker design demonstrate significant listener effects (Mangan, 1961; Markides, 1970; McGarr, 1978; McGarr & Gelfer, 1983; Thomas, 1963), while those using a randomized speaker design (Gulian & Hinds, 1981, present study) show reduced or nonsignificant differences. This would imply that the experienced listener requires exposure to more than a simple sentence to adjust to the talker's idiosyncrasies. A direct comparison between the two experimental designs would be conclusive.

Second, there may be qualitative differences between the two groups of listeners with respect to the way they use articulatory/acoustic information in a carrier phrase and/or sentence context. The significant Listener x Condition x Vowel interaction found here and the lack of a significant listener advantage in the segmented word condition in McGarr's (1978) study suggest differences between the two groups in the way they integrate information within an utterance.

Finally, the listener effect has been noted repeatedly with reference to the perception of words spoken in isolation. It may be that consonant distinctions are more sensitive to the effects of listener experience than judgements of vowel quality.

CRITERIA FOR CORRECT VOWEL IDENTIFICATION

It is useful to understand the relation between acoustics and perception in the speech of the hearing impaired. Training programs which hope to make use of objective measurement devices in speech evaluation and/or self-instruction require objective criteria for acceptable performance. These criteria must be determined in perceptual experiments and then translated into objective confidence limits.

Studies attempting to predict vowel perception from objective data have met with mixed results. Angelocci et al. (1964) report that correctly perceived tokens could not have been predicted from formant values. This conclusion was based on overlap between the F1/F2 vowel spaces of those tokens perceived correctly by at least 75% of the listeners. It should be noted that the formant plots included data from 18 different speakers. Therefore, at least some of the overlap must have resulted from interspeaker differences in, for example, vocal tract length.

Investigations of single talkers (McGarr & Gelfer, 1983; Suonpaa & Aaltonen, 1981) report a fairly predictable relation between formant values and perceptual judgements. The present study corroborates these findings. That is, listeners clearly delineate the non- overlapping portions of individual vowel spaces as acceptable areas. The notion that correct perception can be predicted on the basis of non-overlapping F1/F2 areas is supported by the high correlation between statistical and perceptual clarity. The absolute percent correct scores are not always equivalent, of course, since the predicted classifications consider the normal formant values to be an irrelevant factor.

It is not the case that the absolute values of the formants of 'correct' tokens are constant across talkers or identical to those seen in the normal controls. Osberger et al. (1979) found the same to be true for the F1/F2 ratio in correctly perceived vowels spoken by the hearing impaired. One might ask then, how listeners disambiguate one vowel target from another in a mixed speaker condition (that is, deaf and hearing talkers). Perhaps they alter their criteria for deaf vs. hearing speakers' vowels. The absence of a listener effect speaks against this notion. Furthermore, if this were the case, one would have to predict a significant relationship between correctly identifying the speaker as hearing impaired and recognizing the vowel as that intended by the talker. In the majority of cases, this relation was not significant. For those Talker/Vowel/Condition combinations when it was, the predicted effect was reversed. That is, correctly recognizing the speaker as hearing impaired was associated with an increase in vowel identifiction errors.

Therefore, one must come to the same conclusion regarding the use of steady state cues as that reached with reference to decoding strategy. That is, listeners do not adjust their perceptual criteria with respect to the hearing status of the talker. Kent (1979) developed 'isovowel lines' as a practical method for dealing with intersubject variability in acceptable vowel formants. Isovowel lines are defined as a "linear approximation to formant frequency data for a heterogenous group of" normal subjects plotted in a F1/F2 plane (p.574).

Collection of acoustic and perceptual data from a large sample of hearing-impaired speakers would allow one to delineate areas on the isovowel graphs to be considered the 'range of acceptable formant values'. In a training program using visual, tactile, and/or other computerized speech training devices, this kind of a 'template' would be useful for teacher-directed as well as student-directed therapy.

CONCLUDING REMARKS

What are we to take away from this discussion? It appears that hearing-impaired talkers attempt to differentiate between vowels on the basis of the conventional articulatory variables, but that they are only moderately successful. Token-to-token variability is high. These children would undoubtedly benefit from additional practice and repetition of the motor patterns associated with different speech targets. More work is needed to determine whether other speech contrasts are as variable as vowel contrasts seem to be. Physiological studies are needed so that we can do more than speculate as to the underlying deviant nature of the articulatory gestures used by deaf talkers.

Listeners seem to go about the task of decoding deaf vowels in the 'usual' way. However, it would seem that reduced and/or abnormal coarticulatory patterns in the speech of the deaf must ultimately have effects on the ability of the listener to use this information in interpreting the signal. Vowel perception was apparently resistant to such effects. Future research of this kind on the perception of consonant contrasts may be more revealing.

The results of the present study still leave many questions unanswered regarding the differences between experienced and inexperienced listeners. Although elusive, an understanding of the nature of the experienced listener advantage in the perception of deaf speech may be a key factor in the discovery of some order in an apparently disordered system. If we could describe the unique nature of the signal's unintelligibility, we might be more effective in our attempts to improve the communication skills of hearing-impaired children.

CHAPTER VI

SUMMARY

The results of physical and perceptual studies have led to several theories regarding the deviant nature of speech production in the hearing impaired. An investigation of vowel production in this population provided a forum for testing hypotheses regarding target configurations, as well as those which stress the importance of coarticulation. The following specific questions were addressed:

1. How do hearing-impaired talkers compare to normal hearing controls regarding intrasubject variability in vowel production?
2. Does vowel production in the hearing impaired give evidence for a stable but deviant phonology or for the absence of articulatory/acoustic targets?
3. Do normal hearing listeners use the same acoustic sources of information in determining vowel identity in the speech of the hearing impaired as in the speech of normal hearing subjects?
4. Do experienced and inexperienced listeners differ in ability and/or strategy in identifying vowels in the speech of the deaf?
5. Do experienced and/or inexperienced listeners require dynamic vocal tract information in order to identify a speaker correctly as hearing impaired, and does identification of hearing status interact with phonetic decisions?

Fifteen repetitions of a set of seven vowels (/i,I,ɛ,ɑ,ʌ,ʊ,u/) were spoken by six hearing-impaired and two normal hearing children. Each vowel target was produced in a carrier phrase and fixed consonant context (/bVb/).

F1, F2, f0, and duration measurements were obtained for all tokens. The mean value of each parameter and its variability were determined as a function of vowel target for individual talkers. These estimates served as indicators of articulatory precision and the presence or absence of articulatory/acoustic vowel targets.

All tokens were heard by experienced and inexperienced listeners under conditions providing controlled amounts of vocalic information. That is, each token was heard in three conditions:

1. Sentence Condition (in the context of the carrier phrase)
2. Word Condition (in the /bVb/ syllable segmented from the sentence)
3. Gated Vowel Condition (a 100 ms. steady state portion of the vowel)

In each instance, listeners were asked to identify the test vowel as well as the hearing status of the talker. Their performance as a function of condition allowed for a systematic investigation of the information used in vowel perception and in the identification of the 'deaf voice'.

Token-to-token variability in the four acoustic dimensions considered was greater in the hearing-impaired talkers than in the normal hearing controls. It could not be said, then, that these talkers failed to move their articulators for vowel differentiation. The results of formant analyses, however, do not support the theory of an <u>absence</u> of vowel targets. It was not the case that talkers were random in their employment of articulatory gestures for any given vowel. Instead, these talkers appear to have developed vowel systems based on a reduced number of targets. The framework for one such system was the production of point vowels, while for another it was the preservation of the front/back distinction. In virtually all cases, however, targets were specified by F1 and F2 values. That is, with the exception of a few isolated tense/lax pairs, individual vowel targets were no better specified in an alternative and/or multidimensional acoustic space than in the standard F1/F2 vowel space. There was no evidence to suggest that poor control of a given parameter was compensated for by fine control in another. Thus, there was no support for the notion of a deviant but well-defined phonology.

Perceptual results showed the absence of a calibration effect in vowel perception. That is, listeners did not significantly benefit from information in a carrier phrase, whether the speaker was hearing or hearing-impaired. The consonant context effect was significant. Furthermore, it was significant for hearing and hearing-impaired talkers alike. That is, regardless of the talker, listeners were better able to identify vowels when provided with transitional information than when provided with steady state information alone. Therefore, the results of this study suggest that listeners use the same acoustic sources of information in decoding vowels produced by hearing and hearing-impaired talkers. Formant plots of correctly perceived tokens and correlations between predicted and obtained identification scores confirm the notion that the listener goes about the task of perceiving vowels in a way that is unaffected by the hearing status of the talker.

There was no significant difference between experienced and inexperienced listeners in any of the perceptual tasks. The absence of a listener effect is consistent with the notion that hearing-impaired talkers have not developed deviant but well-defined phonologies. Instead, they operate within systems which can be defined by the normal parameters and are, as such, available to the naive listener.

Results indicate that the ability to identify the talker as hearing-impaired decreases as a function of removing dynamic articulatory information. Performance in the Gated Vowel condition was, however, significantly better than chance. Thus, there is apparently information in the steady state signal which is relevant to this perceptual task.

Appendix A1

Sample Response Form: Sentence

YOU GOT ME THE .....

1.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
2.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
3.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
4.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
5.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
6.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
7.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
8.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
9.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
10.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
11.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
12.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
13.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
14.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
15.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
16.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
17.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
18.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
19.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
20.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
21.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
22.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
23.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI
24.	...beeb	...bib	...behb	...bahb	...buhb	...bub	...boob	H	HI

Appendix A2

Sample Response Form: Word

1.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
2.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
3.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
4.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
5.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
6.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
7.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
8.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
9.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
10.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
11.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
12.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
13.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
14.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
15.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
16.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
17.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
18.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
19.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
20.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
21.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
22.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
23.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI
24.	beeb	bib	behb	bahb	buhb	bub	boob	H	HI

Appendix A3

Sample Response Form: Gated Vowel

1.	ee	i	eh	ah	uh	u	oo	H	HI
2.	ee	i	eh	ah	uh	u	oo	H	HI
3.	ee	i	eh	ah	uh	u	oo	H	HI
4.	ee	i	eh	ah	uh	u	oo	H	HI
5.	ee	i	eh	ah	uh	u	oo	H	HI
6.	ee	i	eh	ah	uh	u	oo	H	HI
7.	ee	i	eh	ah	uh	u	oo	H	HI
8.	ee	i	eh	ah	uh	u	oo	H	HI
9.	ee	i	eh	ah	uh	u	oo	H	HI
11.	ee	i	eh	ah	uh	u	oo	H	HI
12.	ee	i	eh	ah	uh	u	oo	H	HI
13.	ee	i	eh	ah	uh	u	oo	H	HI
14.	ee	i	eh	ah	uh	u	oo	H	HI
15.	ee	i	eh	ah	uh	u	oo	H	HI
16.	ee	i	eh	ah	uh	u	oo	H	HI
17.	ee	i	eh	ah	uh	u	oo	H	HI
18.	ee	i	eh	ah	uh	u	oo	H	HI
19.	ee	i	eh	ah	uh	u	oo	H	HI
20.	ee	i	eh	ah	uh	u	oo	H	HI
21.	ee	i	eh	ah	uh	u	oo	H	HI
22.	ee	i	eh	ah	uh	u	oo	H	HI
23.	ee	i	eh	ah	uh	u	oo	H	HI
24.	ee	i	eh	ah	uh	u	oo	H	HI

Appendix B

INSTRUCTIONS FOR LISTENERS

The purpose of this experiment is to learn more about the way in which hearing-impaired children produce vowels.

Your listening task will be as follows:

You will be listening to tapes of children talking.
Some of the children will be normal hearing and some will be hearing impaired.
On every trial you must decide:

1. which of the seven vowels on your answer sheet best describes the one you've heard
   and...
2. whether the speaker was hearing (H) or hearing impaired (HI).

You will hear 3 different kinds of tapes:

a. On one tape, each trial will consist of the sentence:
   "YOU GOT ME THE ......"
   followed by a "b-vowel-b" nonsense word.
   For example:
   "YOU GOT ME THE beeb.
   You must indicate what vowel you heard in the last word and whether you thought that the speaker was hearing (H) or hearing impaired (HI).
b. On another tape you will hear only a single test word.
   It will again be a "b-vowel-b" nonsense word.
   For example:
   "boob"
   Again, you must indicate what vowel you heard in the nonsense word and whether the speaker was hearing (H) or hearing impaired (HI).
c. Each trial on a third tape will consist of only a vowel sound.
   For example:
   "ee"
   You must indicate what vowel you heard and whether the speaker was hearing (H) or hearing impaired (HI).

The seven vowel choices are shown on you answer sheets using symbols with which you may or may not be familiar.
Here is how you should read them:

(next page)

1. ee/beeb is like the sound in the word "KEYED"
2. i/bib is like the sound in the word "KID"
3. eh/behb is like the sound in the word "KED"
4. ah/bahb is like the sound in the word "COD"
5. uh/buhb is like the sound in the word "COULD"
6. u/bub is like the sound in the word "CUD"
7. oo/boob is like the sound in the word "COOED"

Please review these symbols now to be sure that you know the sound to be associated with each symbol.

For practice, try now to point to the vowel or nonsense word I'm saying:

ee	i	eh	ah	uh	u	oo	H	HI
beeb	bib	behb	bahb	buhb	bub	boob	H	HI

On the test, you will check or circle your answers.

The "H" on your response form is to indicate that you thought the talker was hearing. The "HI" is to indicate that you thought the talker was hearing impaired. In this case you would have marked "H" on every trial recognizing me as a normal hearing talker.

It is important that you answer EVERY trial even if you are only guessing. That is, on every trial you should mark off 2 things:

1. which vowel sound you heard
2. either H or HI

At the beginning of each tape you will hear 14 practice items. These will be followed by 168 test items. You will have 5 seconds in between each trial in which to write your answers. The entire test lasts approximately one hour.

REMEMBER, ON EACH TRIAL YOU MUST:

MARK THE VOWEL SOUND YOU'VE HEARD.

MARK EITHER H OR HI.

DO NOT LEAVE ANY BLANKS. IF YOU ARE NOT SURE OF AN ANSWER...PLEASE GUESS.

Are there any questions?

Appendix C1

Normalized Values of f0 as a Function of Vowel Target

	/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/
D1	1.22	1.22	1.24	1.0	1.14	1.25	1.18
D2	1.04	1.03	1.02	1.0	1.01	1.02	1.02
D3	1.11	1.12	1.12	1.0	1.08	1.11	1.11
D4	1.18	1.17	1.14	1.0	1.10	1.07	1.17
D5	1.18	1.13	1.13	1.0	1.07	1.19	1.25
D6	1.16	1.04	1.07	1.0	1.03	1.05	1.16
H1	1.04	1.05	1.05	1.0	1.06	1.09	1.08
H2	1.07	1.06	1.04	1.0	1.03	1.09	1.10
A	1.15	1.15	1.13	1.0	1.09	1.09	1.14
B	1.13	1.13	1.13	1.0	1.08	1.14	1.10
C	1.17	1.09	1.10	1.0	1.05	1.12	1.21
H	1.06	1.06	1.05	1.0	1.05	1.09	1.09

Appendix C2

Normalized Values of Duration as a Function of Vowel Target

	/i/	/I/	/ɛ/	/ɑ/	/ʌ/	/ʊ/	/u/
D1	.66	.78	.75	1.0	.50	.55	1.11
D2	.77	.71	.71	1.0	.66	.65	.79
D3	.81	.52	.58	1.0	.49	.50	.87
D4	.92	.72	.66	1.0	.77	.72	.96
D5	.74	.61	.73	1.0	.61	.70	.69
D6	.68	.54	.60	1.0	.49	.50	.61
H1	.88	.57	.64	1.0	.67	.60	.90
H2	.82	.62	.71	1.0	.68	.61	.84
A	.87	.62	.65	1.0	.63	.61	.92
B	.72	.75	.73	1.0	.58	.60	.95
C	.71	.58	.67	1.0	.55	.60	.65
H	.85	.60	.68	1.0	.68	.61	.87

Appendix D

Percent Correct Vowel Recognition as a Function of Vowel for Individual Talkers: Sentence Condition

	D1	D2	D3	D4	D5	D6	H1	H2
/i/	56.7	57.9	32.7	92.9	65.0	57.9	99.3	95.3
/I/	61.4	54.6	92.5	95.4	12.5	47.5	99.6	100.0
/ɛ/	5.8	24.0	3.3	11.9	3.6	17.8	87.0	98.5
/ɑ/	95.3	93.1	98.0	88.7	81.7	92.5	90.7	94.2
/ʌ/	36.7	10.0	14.2	52.3	9.2	40.0	93.8	91.9
/ʊ/	16.3	21.0	52.9	17.1	11.3	40.0	60.0	33.9
/u/	90.3	76.5	82.3	33.6	31.5	62.5	98.9	96.0

Appendix E

Percent Correct Vowel Recognition as a Function of Vowel for Individual Talkers: Word Condition

	D1	D2	D3	D4	D5	D6	H1	H2
/i/	77.1	69.3	27.3	95.7	82.0	66.7	96.1	94.3
/I/	32.5	47.3	91.0	91.3	5.0	40.0	96.3	99.3
/ɛ/	4.2	35.0	5.6	26.5	5.0	17.2	89.0	97.7
/ɑ/	97.0	92.3	96.3	94.0	84.7	90.4	80.0	93.1
/ʌ/	40.0	12.0	21.7	53.6	4.2	50.7	92.9	95.6
/ʊ/	19.4	33.0	56.4	15.4	5.0	32.5	48.3	26.1
/u/	83.3	50.8	60.7	26.1	19.0	49.0	94.3	97.0

Appendix F

Percent Correct Vowel Recognition as a Function of Vowel for Individual Talkers: Gated Vowel Condition

	D1	D2	D3	D4	D5	D6	H1	H2
/i/	75.0	45.0	24.0	97.1	73.0	60.4	97.9	93.3
/I/	32.1	36.9	55.0	81.7	12.5	5.0	90.8	92.1
/ɛ/	4.2	45.7	8.9	28.8	10.7	15.6	89.5	97.3
/ɑ/	94.7	67.7	83.5	74.7	69.7	78.2	59.7	92.3
/ʌ/	26.7	32.5	19.2	30.0	0.00	38.3	73.3	61.3
/ʊ/	22.5	52.0	64.3	12.5	10.0	37.5	51.7	18.3
/u/	72.7	15.0	58.7	27.9	19.5	24.5	92.1	87.7

Appendix G

Percent Correct Vowel Recognition as a Function of Vowel and Condition: Hearing vs. Hearing-Impaired Talkers

	NORMALS (n=2)				DEAF (n=6)			
	S	W	V	Mean	S	W	V	Mean
/i/	97.2	95.1	95.5	96.0	60.0	68.4	61.1	63.2
/I/	99.7	97.4	91.3	96.1	70.9	59.7	47.1	59.2
/ɛ/	93.5	93.9	93.9	93.8	12.3	17.8	21.6	17.3
/ɑ/	92.3	86.1	74.8	84.4	91.5	92.5	78.3	87.4
/ʌ/	93.0	94.0	68.5	85.2	28.3	31.4	25.8	28.5
/ʊ/	40.4	31.7	26.7	32.9	23.1	24.9	31.0	26.3
/u/	97.4	95.7	89.8	94.3	64.9	50.2	38.9	51.3

REFERENCES

Ainsworth, W. A. (1972). Duration as a cue in the recognition of synthetic vowels. Journal of the Acoustical Society of America, 51, 648-651.

Ainsworth, W. A. (1974). The influence of precursive sequences on the perception of synthesized vowels. Language and Speech, 17, 103-109.

Angelocci, A., Kopp, G., & Holbrook, A. (1964). The vowel formants of deaf and normal hearing eleven to fourteen year old boys. Journal of Speech and Hearing Disorders, 29, 156-170.

Assmann, P. F., Nearey, T. M., & Hogan, J. T. (1982). Vowel identification: Orthographic, perceptual, and acoustic aspects. Journal of the Acoustical Society of America, 71, 975-989.

Bailey, P. J., & Summerfield, Q. (1980). Information in speech: Observations on the perception of [s]-stop clusters. Journal of Experimental Psychology: Human Perception and Performance, 6, 536-563.

Bell-Berti, F., & Harris, K. S. (1976). Some aspects of coarticulation. Haskins Laboratories Status Report on Speech Research, SR-45/46, 197-204.

Best, C. T., Morrongiello, B., & Robson, R. (1981). Perceptual equivalence of acoustic cues in speech and nonspeech perception. Perception & Psychophysics, 29, 191-211.

Bond, (1975). Identification of vowels excerpted from context. Journal of the Acoustical Society of America, 57, 524.

Bond, (1976). Identification of vowels excerpted from /l/ and /r/ contexts. Journal of the Acoustical Society of America, 60, 906-910.

Boone, D. (1966). Modification of the voices of deaf children. Volta Review, 68, 686-692.

Boothroyd, A. (1980). Evaluation of speech. SARP, 34.

Boothroyd, A., & Gorzycki (1977). Sibilant articulation in hearing impaired children. ASHA Meeting.

Boothroyd, A., Nickerson, R., & Stevens, K. (1974). Temporal patterning in the speech of the deaf. (SARP Report #7). Northampton: Clarke School for the Deaf.

Brannon, J. (1966). The speech production and spoken language of the deaf. Language and Speech, 9, 127-136.

Broad, D. J. (1976). Toward defining acoustic phonetic equivalence for vowels. Phonetica, 33, 401-424.

Brownlee K.A., (1965). Statistical Theory and Methodology in Science and Engineering. New York: John Wiley and Sons.

Bush, M. (1981). Vowel articulation and laryngeal control in the speech of the deaf. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

Calvert, D. R. (1961). Some acoustic characteristics of the speech of profoundly deaf individuals. Unpublished doctoral dissertation, Stanford University.

Crouter (1963) A cinefluorographic comparison of selected vowels spoken by deaf and hearing subjects. Unpublished master's thesis, University of Kansas.

Dechovitz, D. (1977). Information conveyed by vowels: A confirmation. Haskins Laboratories Status Report on Speech Research, SR-51/52, 213-219.

Diehl, R. L., McCusker, S. B., & Chapman, L. S. (1981). Perceiving vowels in isolation and in consonantal context. Journal of the Acoustical Society of America, 68, 239-248.

Dorman, M. F., Raphael, L. J., & Liberman, A. M. (1979). Some experiments on the sound of silence in phonetic perception. Journal of the Acoustical Society of America, 65, 1518-1532.

Eguchi, & Hirsh, I. (1969). Development of speech sounds in children. Acta Oto-laryngia, Spp. 257.

Fairbanks, G., & Grubb, P. A. (1961). A psychophysical investigation of vowel formants. Journal of Speech and Hearing Research, 4, 203-219.

Fitch, H. L., Halwes, T., Erickson, D. M., & Liberman, A. M. (1980). Perceptual equivalence of two acoustic cues for stop consonant manner. Perception & Psychophysics, 27, 343-350.

Forner, L., & Hixon, T. (1977). Respiratory kinematics in profoundly hearing-impaired speakers. Journal of Speech and Hearing Research, 20, 373-408.

Fowler, C. A. (1980). Coarticulation and theories of extrinsic timing control. Journal of Phonetics, 8, 113-133.

Fowler, C. A. (1981). Perception and production of coarticulation among stressed and unstressed vowels. Journal of Speech and Hearing Research, 46, 127-139.

Fowler, C. A., & Shankweiler, D. P. (1978). Identification of vowels in speech and nonspeech contexts. Journal of the Acoustical Society of America, 63, S4. (Abstract)

Fujimura, O., & Ochiai, K. (1963). Vowel identification and phonetic contexts. *Journal of the Acoustical Society of America*, *35*, 1889. (Abstract)

Gerratt B., (1983). Formant frequency fluctuation as an index of motor steadiness in the vocal tract. *Journal of Speech and Hearing Research* *26*, 297-304.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton-Mifflin.

Gold, T. (1978). *Speech and hearing skills: A comparison between hard-of-hearing and deaf children*. Unpublished doctoral dissertation, City University of New York.

Gottfried, T. L., & Strange, W. (1980). Identification of coarticulated vowels. *Journal of the Acoustical Society of America*, *68*, 1626-1635.

Gulian, & Hinds (1981). Response bias in the perception of deaf speech by experienced and naive listeners. *British Journal of Audiology*, *15*, 165-171.

Haggard, (1971). Perceptual processing of coarticulation: A case study of /l/. *Journal of Phonetics*, *2*, 117-123.

Harris, K. S. (1978). Physiological aspects of speech production. In J. F. Kavanaugh & W. Strange (Eds.), *Speech and language in the laboratory, school and clinic*. Cambridge, MA: MIT Press.

Harris, K. S., & McGarr, N. S. (1980). Relationships between speech perception and speech production in normal hearing and hearing impaired subjects: Implications for training speech of the hearing impaired. In J. D. Subtelny (Ed.), *Speech assessment and speech improvement for the hearing impaired* (pp. 316-337). Washington, DC: A. G. Bell Association.

Horwich (1977). *An examination of the vowels of deaf girls.* S.B. Thesis Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

Hudgins, C. V., & Numbers, F. C. (1942). An investigation of the intelligibility of the speech of the deaf. *Genetic Psychology Monographs*, *25*, 289-392.

Howell, P. (1981). Identification of vowels in and out of context. *Journal of the Acoustical Society of America*, 1256-1260.

Huntington, D., Harris, K. S., & Sholes, G. (1968). An electromyographic study of consonant articulation in hearing-impaired and normal speakers. *Journal of Speech and Hearing Research*, *11*, 147-158.

Irvin, B., & Wilson, L. (1973). The voiced-unvoiced distinction in deaf speech. *American Annals of the Deaf*, *118*, 43-45.

John, J. E. J., & Howarth, J. N. (1965). The effect of time distortions on the intelligibility of deaf children's speech. Language and Speech, 8, 127-134.

Joos, M. (1948). Acoustic phonetics. Language Monograph 23, Vol. 24(Suppl.)

Kent, R. (1978). Imitation of syntesized vowels by pre-school children. Joural of the Acoustical Society of America, 63, 1193-1198.

Kent, R. (1979). Isovowel lines for the evaluation of vowel formant structure in speech disorders. Journal of Speech and Hearing Disorders, 44, 513-521.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. Journal of Phonetics, 3, 129-140.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. Journal of the Acoustical Society of America, 59, 1208-1221.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. Journal of the Acoustical Society of America, 29, 98-104.

Lehiste, I., & Meltzer, D. (1973). Vowel and speaker identification in natural and synthetic speech. Language and Speech, 16, 356-364.

Levitt, H., & Stromberg, H. (1972). Digital processing of the speech of deaf children. ACM Annual Conference Proceedings, Boston.

Levitt, H., & Stromberg, H. (1983). Segmental characteristics of the speech of hearing impaired children: Factors affecting intelligibility. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), Speech of the hearing impaired: Research, training, and personnel preparation. Baltimore: University Park Press.

Levitt, H., Stromberg, H., Smith, C., & Gold, T. (1980). The structure of segmental errors in the speech of deaf children. Journal of Communication Disorders, 13, 419-442.

Liberman, A. M. (1982). On finding that speech is special. American Psychologist, 37, 148-167.

Lieberman, P. (1980). On the development of of vowel production in young children. Child phonology, Vol. I: Production. New York: Academic Press.

Lindblom, B. (1962). Accuracy and limitations of sonograph measurements. Proceedings of the 4th International Congress of Phonetic Sciences, Helsinki; 1961, The Hague.

Lindblom, B. E. F., & Studdert-Kennedy, M. (1967). On the role of formant-transitions in vowel recognition. Journal of the Acoustical Society of America, 42, 830-843.

Ling, D. (1976). Speech and the hearing impaired child: Theory and practice. Washington, DC: A. G. Bell.

Macchi, M. J. (1980). Identification of vowels spoken in isolation and in consonantal context. Journal of the Acoustical Society of America, 68, 1636-1642.

Malhalanobis P.C., (1936). On the generalized distance in statistics. Proceedings of the National Institute of Science, India, 12, 49-55.

Mangan, K. (1961). Speech improvement through articulation testing. American Annals of the Deaf, 106, 391-396.

Markides, A. (1970). The speech of deaf and partially hearing children with special reference to factors affecting intelligibility. British Journal of Disorders of Communication, 5, 126-140.

Martony, J. (1968). On the correction of the voice pitch level for severely hard of hearing subjects. American Annals of the Deaf, 113, 195-202.

McGarr, N. S. (1978). The differences between experienced and inexperienced listeners in understanding the speech of the deaf. Unpublished doctoral dissertation, The City University of New York.

McGarr, N. S., & Osberger, M. J. (1978). Pitch deviancy and the intelligibility of deaf children's speech. Journal of Communication Disorders, 11, 237-247.

McGarr, N. S., & Gelfer, C. (1983). Simultaneous measurements of vowels produced by a hearing-impaired speaker. Language and Speech.

McGarr, N. S., & Harris, K. S. (1893). Articulatory control in a deaf speaker. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), Speech of the hearing impaired: Research, training, and personnel preparation. Baltimore: University Park Press.

McGarr, N., et al., (in press). Sytematic Speech Training Procedures and Sensory Aids. C.U.N.Y. Research Grant #142-40-2721

McGarr, N. S., & Lofqvist, A. (1982). Obstruent production in hearing-impaired speakers: Interarticulator timing and acoustics. Journal of the Acoustical Society of America, 72, 34-42.

Metz, D. D. (1980). Morphological boundaries and the timing of stressed vowels produced by profoundly hearing impaired adults. Journal of Phonetics, 8, 63-68.

Monsen, R. B. (1974). Durational aspects of vowel production in the speech of deaf children. Journal of Speech and Hearing Research, 17, 386-398.

Monsen, R. B. (1976a). Normal and reduced-phonological space: The production of English vowels by deaf adolescents. Journal of Phonetics, 4, 189-198.

Monsen, R. B. (1976b). Second formant transitions of selected consonant-vowel combinations in the speech of deaf and normal hearing children. Journal of Speech and Hearing Research, 19, 279-289.

Monsen, R. B. (1976c). A taxonomic study of diphthong production in the speech of deaf children. In S. K. Hirsh, D. H. Eldredge, I. S. Hirsh, & S. R. Silverman (Eds.), Hearing and Davis: Essays honoring Hallowell Davis (pp. 281-290). St. Louis, MO: Washington University Press.

Monsen, R. B. (1976d). The production of English stop consonants in the speech of deaf children. Journal of Phonetics, 4, 29-42.

Monsen, R. B. (1978). Towards measuring how well hearing-impaired children speak. Journal of Speech and Hearing Research, 21, 197-219.

Monsen, R., & Engebretson, M. (1983). The accuracy of formant frequency measurements: a comparison of spectrographic analysis and linear prediction. Journal of Speech and Hearing Research, 26. 89-97.

Nickerson, R. (1973). Computerized speech training aids for the deaf. Bolt, Beranek and Newman, Inc. Report No. 2633.

Nickerson, R. (1975). Characteristics of the speech of deaf persons. Volta Review, 342-362.

Nickerson, R., & Stevens, K. N. (1980). Approaches to the study of the relationship between intelligibility and the physical properties of speech. In J. Subtelny (Ed.), Speech assessment and speech improvement for the hearing impaired. Washington, D.C.: A. G. Bell Association for the Deaf.

Nickerson, R. S., Stevens, K. N., Boothroyd, A., & Rollins, A. (1974). Some observations on timing in the speech of deaf and hearing speakers. (Report No. 2905). Cambridge, MA: Bolt, Beranek and Newman, Inc.

Nober, H. (1967). Articulation of the deaf. Exceptional Children, 33, 611-621.

Ochiai, K., & Fujimura, O. (1971). Vowel identification and phonetic contexts. Reports from the Electrocommunications (University of Tokyo), 22, 103-111.

Ohde, R. N., & Sharf, D. J. (1977). Order effect of acoustic segments of VC and CV syllables on stop and vowel identification. Journal of Speech and Hearing Research, 20, 543-554.

Ohman, S. E. G. (1966). Coarticulation in VCV utterances: Spectrographic measurements. Journal of the Acoustical Society of America, 39, 151-168.

Osberger, M. J., & Levitt, H. (1979). The effect of timing errors on the intelligibility of deaf children's speech. <u>Journal of the Acoustical Society of America</u>, <u>66</u>, 1316-1324.

Osberger, M. J., Levitt, H., & Slosberg, R. (1979). Acoustic characteristics of correctly produced vowels. <u>Journal of the Acoustical Society of America</u>, <u>66</u>, S13. (Abstract)

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. <u>Journal of the Acoustical Society of America</u>, <u>24</u>, 175-184.

Pisoni, D. (1976). Vowel mimicry and the quantal theory of speech: A first report. <u>Research on Speech Perception Progress Report</u> (Indiana University), <u>3</u>.

Pisoni, D., Carrell, T. D., & Simnick, S. S. (1979). Does a listener need to recover the dynamic vocal tract gestures of a talker to recognize his vowels? <u>Journal of the Acoustical Society of America</u>, <u>65</u>, 5.

Potter, R. K., & Steinberg, J. C. (1950). Toward the specification of speech. <u>Journal of the Acoustical Society of America</u>, <u>22</u>, 807-820.

Reilly, A. P. (1979). <u>Syllabic nucleus duration in the speech of hearing and deaf children</u>. Unpublished doctoral dissertation, The City University of New York.

Rothman, H. (1976). A spectrographic investigation of consonant-vowel transitions in the speech of deaf adults. <u>Journal of Phonetics</u>, <u>4</u>, 129-136.

Rothman, H. (1977). An electromyographic investigation of articulation and phonation patterns in the speech of deaf adults. <u>Journal of Phonetics</u>, <u>5</u>.

Shankweiler, D. P., Strange, W., & Verbrugge, R. R. (1977). Speech and the problem of perceptual constancy. In R. E. Shaw & J. D. Branford (Eds.), <u>Perceiving, acting, and knowing: Toward an ecological psychology</u>. Hillsdale NJ: Erlbaum.

Shankweiler, D., Verbrugge, R. R., & Studdert-Kennedy, M. (1978). Insufficiency of the target for vowel perception. <u>Haskins Laboratories Status Report on Speech Research</u>, <u>SR-55/56</u>, 103-111.

Simon,C., (1983). Relationship between production and perception of speech: some developmental issues. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), <u>Speech of the hearing impaired: Research, training and personnel preparation</u>. Baltimore, MD: University Park Press.

Smith, C. R. (1975). Residual hearing and speech production of deaf children. <u>Journal of Speech and Hearing Research</u>, <u>18</u>, 795-811.

Stein, D. (1980). <u>A study of articulatory characteristics of deaf talkers</u>. Unpublished doctoral dissertation, University of Iowa.

Stevens, K. N. (1973). The quantal nature of speech: Evidence from articulatory-acoustic data. In E. E. David & P. B. Denes (Eds.), Human communication: A unified view. New York: McGraw Hill.

Stevens, K. (1977). Physics of laryngeal behavior and larynx modes. Phonetica, 34, 264-279.

Stevens, K. N., & House, A. S. (1955). Development of a quantitative description of vowel articulation. Journal of the Acoustical Society of America, 27, 484-493.

Stevens, K. N., & House, A. S. (1963). Perturbation of vowel articulations by consonantal context: An acoustical study. Journal of Speech and Hearing Research, 6, 111-128.

Stevens, K., Kalikow, D.N., & Willemain, T.R. (1975). A miniature accelerometer for detecting glottal waveforms and nasalization. Journal of Speech and Hearing Research, 18, 594-599.

Stevens, K. N., Nickerson, R. S., & Rollins, A. M. (1983). Suprasegmental and postural aspects of speech production and their effect on articulatory skills and intelligibility. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), Speech of the hearing impaired: Research, training and personnel preparation. Baltimore, MD: University Park Press.

Strange, W., Edman, T. R., & Jenkins, J. J. (1979). Acoustic and phonological factors in vowel identification. Journal of Experimental Psychology: Human Perception and Performance, 5, 643-656.

Strange, W., & Gottfried, T. L. (1980). Task variables in the study of vowel perception. Journal of the Acoustical Society of America, 68, 1622-1625.

Strange, W., Jenkins, J. J., & Edman, T. R. (1977). Identification of vowels in 'vowel-less' syllables. Journal of the Acoustical Society of America, 61.

Strange, W., Jenkins, J. J., & Edman, T. R. (1978). Dynamic information specifies vowel identity. Journal of the Acoustical Society of America, 63. (Abstract)

Strange, W., Verbrugge, R., Shankweiler, D. P., & Edman, T. R. (1976). Consonantal environment specifies vowel identity. Journal of the Acoustical Society of America, 60, 213-224.

Summerfield, Q., & Haggard, M. (1977). On the dissociation of spectral and temporal cues to the voicing distinction in initial stop consonants. Journal of the Acoustical Society of America, 62, 435-448.

Suonpaa, J., & Aaltonen, G. (1981). Intelligibility of vowels in words uttered by profoundly hearing-impaired children. Journal of Phonetics, 9, 445-450.

Suonpaa, J., & Aaltonen, G. (1982). The validity of real time spectrum analysis of Finnish vowel formants in profoundly hearing-impaired children. *Scandinavian Audiology*, *11*, 43-48.

Thomas, W. (1963). Intelligibility of the speech of deaf children. *Proceedings of the International Congress on the Education of the Deaf*, Document No. 196.

Tuller,B., & Harris,K., (1980) Temporal models of interarticulator programming. *Journal of the Acoustical Society of America*, *67*, Abstract

Tuller, B., Kelso, J. A. S., & Harris, K. S. (in press). Converging evidence for the role of relative timing in speech. *Journal of Experimental Psychology: Human Perception and Performance*.

Verbrugge, R. R., & Shankweiler, D. P. (1977). Prosodic information for vowel identity. *Journal of the Acoustical Society of America*, *61*, S39. (Abstract)

Verbrugge, R. R., & Isenberg, D. (1978). Syllable timing and vowel perception. *Journal of the Acoustical Society of America*, *63*, S4. (Abstract)

Verbrugge, R. R., Strange, W., Shankweiler, D. P., & Edman, T. R. (1976). What information enables a listener to map a talker's vowel space? *Journal of the Acoustical Society of America*, *60*, 198-212.

Whitehead, R. (1983). Some respiratory and aerodynamic patterns in the speech of the hearing impaired. In I. Hochberg, H. Levitt, & M. J. Osberger (Eds.), *Speech of the hearing impaired: Research, training, and personnel preparation*. Baltimore, MD: University Park Press.

Whitehead, R., & Jones, (1976). Influence of consonant environment on duration of vowels produced by normal hearing and hearing impaired adult speakers. *Journal of the Acoustical Society of America*, *60*, 513-514.

Whitehead, R., & Jones, (1978). The effect of vowel environment on duration of consonants produced by normal hearing; hearing impaired and deaf adult speakers. *Journal of Phonetics*, 77-81.

Whitehead, R., & Barefoot, S. (1980). Some aerodynamic characteristics of plosive consonants produced by hearing-impaired speakers. *American Annals of the Deaf*, *125*, 366-373.