VOWEL ARTICULATION AND LARYNGEAL CONTROL IN THE SPEECH OF THE DEAF

by

Marcia Ann Bush

B.S., Pennsylvania State University
(1974)

S.M., Massachusetts Institute of Technology
(1977)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

July, 1981

Signature of Author: ______________________________
Department of Electrical Engineering and Computer Science
July 30, 1981

Certified by: ______________________________
Kenneth N. Stevens

Accepted by: ______________________________
Arthur C. Smith
Chairman, Departmental Committee on Graduate Students

VOWEL ARTICULATION AND LARYNGEAL CONTROL IN THE SPEECH OF THE DEAF

by Marcia Ann Bush

Submitted to the Department of Electrical Engineering and Computer Science on July 30, 1981, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

The aim of this thesis was to document and to explain some of the linguistically inappropriate variations in voice fundamental frequency (F0) which characterize the utterances of deaf speakers. The specific objectives of the thesis were twofold: 1) to examine more carefully the relationship between the vowel-to-vowel variations in F0 produced by deaf speakers and their proficiency at vowel articulation; and 2) to explore mechanisms which might account for the exaggerated vowel-related variations in F0 observed for many deaf speakers. These two objectives were pursued in the thesis through an analysis of F0, formant and spectral data for groups of deaf and hearing boys and girls and through a comparison of these data with intelligibility scores for the deaf speakers. The collection and interpretation of the data were guided by a consideration of mechanisms proposed to account for interactions between vowel height and F0 in the speech of normal-hearing individuals.

For the majority of deaf speakers in the study (sixteen of twenty), a close relationship was observed between vowel-related variability in F0 and articulatory skill. In general, greater F0 variability and higher intelligibility scores were observed for speakers who produced a relatively wide range of vowel sounds (i.e., of first- and second-formant frequency values) than for speakers whose articulatory capabilities were more limited. Exaggerated vowel-to-vowel variations in F0 were produced by deaf speakers who maintained a mean F0 which was somewhat higher than normal, and who were capable of articulatory configurations appropriate to high vowels. The amount of F0 variability used by these speakers was determined primarily by an excessively high F0 for the high vowels /i,I,u/ relative to F0 for /ɑ/. Smaller vowel-to-vowel variations in F0 were produced by deaf speakers whose mean F0 was comparable to normal and by speakers whose articulatory skills, particularly with respect to the production of high vowels, were poor.

The segmental variations in F0 produced by these speakers appeared to be best explained by an extension of a mechanism proposed by Honda to account for normal vowel-related variations in F0 [Hon81]. Honda's mechanism assumes that shifting the tongue root forward for the production of high vowels will also cause the hyoid bone to move forward and to tilt the thyroid cartilage anteriorly, resulting in an increased longitudinal tension on the vocal folds and, thus, an increase in F0. On the basis of the non-linear nature of the stress-strain relationship for vocal-fold tissue, it was argued that such increases in vocal-fold tension may be somewhat greater in magnitude when the tension on the vocal folds (and, thus, mean F0) is already relatively high, leading to somewhat larger increases in F0 during the articulation of high vowels. This argument appears to be consistent with F0 data reported in the literature for hearing speakers as well as with the interactions among F0 variability, mean F0 and articulatory proficiency described above.

The exaggerated vowel-related variations in F0 produced by a smaller group of deaf speakers (four girls) were more difficult to explain on the basis of this mechanism. While all four speakers produced the vowel /ɑ/ with an F0 comparable to normal and the vowels /i/ and /u/ with excessively high F0, they differed considerably among themselves with respect to mean F0 and articulatory skill. Each of the girls was, however, approximately fourteen years old, suggesting that age-related factors (e.g., adolescent voice change or similarities of speech training) may have contributed to their problems with vowel-related F0 control.

The close interrelationship between vowel production and F0 observed for the deaf boys and girls in the study indicates the need for teachers of the deaf to monitor the adequacy of segmental articulation during attempts to train more stable fundamental-frequency control.

Thesis Supervisor: Professor Kenneth N. Stevens
Title: LeBel Professor of Electrical Engineering

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to the following:

Kenneth Stevens, for his advice and friendship;

Victor Zue and Louis Braida, for serving as readers;

Sarah Hawkins, Martha Laferriere and Lise Menn, for their painstaking transcriptions;

Ursula Goldstein, William Henke and Keith North, for computer programs and technical assistance;

Anne Allen and the staff of the M.I.T. High School Studies Program, for their help in recruiting speakers;

The faculty and staff of the Clarke School for the Deaf, particularly Arthur Boothroyd, Beth Bellinger and Patricia Archambault, for their time and generosity;

The Whitaker Health Sciences Fund, for financial support as a graduate fellow;

My mother, Marie Bushinski, and my husband, Gary Kopec, for their faith and patience;

The boys and girls who served as speakers, for making this study possible.

TABLE OF CONTENTS

LIST OF FIGURES

Figure | Page

Figure Page

Figure Page

LIST OF TABLES

CHAPTER 1

INTRODUCTION AND BACKGROUND

The general objective of this thesis is to gain an understanding of the relationship between vowel articulation and laryngeal control in the speech of profoundly hearing-impaired boys and girls. The approach is to carry out an acoustic analysis of their speech and to draw inferences about the speech mechanisms involved on the basis of the acoustic data. This line of research is motivated by two observations: 1) large deviations in F0 appear to be related to the production of <u>specific</u> vowel sounds in the speech of some hearing-impaired individuals; and 2) comparable but smaller interactions between vowel articulation and F0 also occur in the utterances of normal-hearing speakers. A major question addressed in the thesis is whether the effects of such normal interactions on F0 are exaggerated for some hearing-impaired speakers due to inappropriate laryngeal postures or to extreme articulatory maneuvers used in vowel and consonant production.

The analyses performed as part of this study are motivated more fully in the following sections of this chapter. First, literature relating to the production of exaggerated deviations in F0 by hearing-impaired speakers is reviewed (Section 1.1), as are mechanisms proposed to account for normal interactions between vowel articulation and F0 (Section 1.2). Next, a number of hypotheses are made about how certain aspects of the speech of the hearing-impaired (e.g., excessive jaw movement) might influence the operation of these mechanisms, resulting in larger-than-normal vowel-related changes in F0 (Section 1.3). Finally, the specific research objectives of the thesis are defined and an overview of the remaining chapters of the thesis is provided (Section 1.4).

Before these tasks are begun, however, a note on terminology is appropriate. In the remainder of this thesis, the terms "deaf" and "hearing" will be used in referring to profoundly hearing-impaired and normal-hearing speakers respectively. For the speakers participating in the present study, profound hearing impairment implies an average loss of at least 80 db in the "speech-frequency" range (i.e., 500 Hz, 1KHz and 2Kz) [Mea80]. In the literature review and discussions which follow, the term "hearing-impaired" is used to refer to speakers whose losses are known to be less severe.

1.1 Excessive F0 Variability in the Speech of the Deaf

The production of large and/or incorrectly located variations in voice fundamental frequency (i.e., relative to those produced by hearing speakers) characterizes the speech of many deaf individuals. This problem is often referred to as one of "erratic" pitch, in that there is usually no apparent linguistic or intonational motive (e.g., stress assignment, interrogation) for the deaf speaker's use of exaggerated F0 change [Nic75]. There is evidence to suggest, however, that large deviations in F0 are, in fact, produced systematically by some deaf speakers. For example, several investigators have reported that vowel-to-vowel variations in F0 tend to be exaggerated in the speech of the deaf and that certain vowels (most notably the high vowels /i/ and /u/) are often produced with excessively high F0 [Hor77]; [Mar77];[Mar68];[AngKop64].

The factors which underlie such vowel-related deviations in F0 are not well understood, although a number of explanations have been proposed. Angelocci, Kopp and Holbrook [AngKop64], for example, measured F0 as a function of vowel height for groups of deaf and hearing boys and found that, on average, F0 varied more from vowel to vowel for the deaf speakers. The opposite was true, however for the vowel formant frequencies (i.e., F1 and F2 spanned a wider range for the hearing speakers), leading the authors to suggest that the deaf boys might be using excessive laryngeal variability rather than accurate placement of the articulators to distinguish one vowel sound from another.

A similar conclusion was reached by Horwich [Hor77] who compared F0 and formant-frequency data for a group of deaf girls with those reported by Peterson and Barney [PetBar52] for hearing men, women and children. Calculations on her data show, however, that large differences existed in the range of cross-vowel fundamental and formant frequencies used by individual deaf speakers. For example, some girls who produced large vowel-to-vowel variations in F0 also produced a relatively wide range of F1 and F2 values, while other girls used relatively little variation in F0, F1 or F2. Such calculations suggest that statements about a trade-off between articulatory proficiency and F0 control based solely on averaged data might be misleading for many deaf speakers.

The tendency for deaf speakers to produce large vowel-related changes in F0 has been noted by a number of other investigators [NicSte77];[Lin76];[Nic75];[Mar68]. However, in contrast to Angelocci, et.al. and to Horwich, these investigators have proposed that excessive F0 variability might be an inadvertent consequence of (rather than a substitute for) the articulatory maneuvers used in vowel production. Martony [Mar68], for example, found that many deaf speakers produced the close Swedish vowels /i,y,ø/ with inappropriately high F0 even when the articulation of these vowels was judged to be quite good. He reasoned that the supraglottal muscles involved in the articulation (perhaps "extreme" articulation) of some vowels and consonants might also influence the tension and position of the larynx, thus leading to exaggerated changes in F0. Ling [Lin76] has attributed the high F0 often associated with the production of the high front vowel /i/ to the inability of some deaf speakers to control fundamental frequency independently of tongue and jaw position, and Stevens [Ste-pc] has suggested that this interdependence between articulation and F0 might be enhanced by an inappropriate (e.g., tense) laryngeal posture. In no case, however, have objective data been collected to specifically test these various hypotheses.

The effects of excessive (vowel-related) F0 variability on speech intelligibility are also not well understood. For example, while a number of authors have reported high negative correlations between intelligibility scores and listener judgements of pitch "breaks" or intermittent phonation [StrLev79];[McGOsb78];[ParLev78], such severe problems with phonatory control appear to occur too infrequently to account for most cases of vowel-related erratic pitch. Stromberg and Levitt [StrLev79] obtained judgements of "inappropriate variability" and "excessive variability" of intonation for a larger proportion of deaf speakers (i.e., 31.6% and 13.6% of forty speakers respectively), but they implied that correlations between these judgements and intelligibility were not very high. (It should be noted that the extent to which such variations in F0 depended upon language deficienies, as opposed to segmental factors such as vowel height, was not considered.) Judgements of "insufficient variability" of intonation, however, which were obtained for 36.7% of the speakers, showed a large negative correlation with intelligibility scores.

Studies comparing objective measurements of F0 variability with speech intelligibility are few in number. Monsen [Mon78] reported very low correlations between intelligibility and measures of mean F0 and F0 range for a group of thirty-seven hearing-impaired children, but noted that his averaged data were probably insensitive to individual differences in the problems occurring among his speakers as well as to the direction of F0 variations (appropriate vs. inappropriate) which these speakers produced. In contrast, Nickerson, et.al. [NicSte79b] reported a high correlation between F0 range and average intelligibility for a group of fifteen deaf boys and girls, most of whom had received training with a visible F0 display. Within this group of speakers, low intelligibility was typically associated with monotone speech, and, somewhat surprisingly, a positive correlation was found between average intelligibility and the amount of <u>deviation</u> from an idealized (sentential) F0 contour. Noting that intelligibility scores for unfiltered speech are probably most reflective of segmental or articulatory skills

[LevSmi74], Nickerson, et.al. reasoned that the F0 deviations produced by their speakers might be associated with the (correct) articulation of specific vowels [Mar68];[Lin76] and, hence, be compatible with good speech intelligibility. While not conclusive, these latter observations lend some support to the hypothesis that articulation and F0 control are closely interdependent for some deaf speakers.

1.2 Vowel Articulation and F0 Control for Hearing Speakers

Systematic variations in voice fundamental frequency as a function of vowel height can also be observed in the speech of hearing individuals. Again, numerous studies have shown that vowels articulated with the tongue body high in the mouth, such as /i/ or /u/, tend, on average, to be spoken with a somewhat higher F0 than vowels articulated with the tongue body lower in the mouth, such as /ɑ/ or /æ/ [Pet76]; [LehPet61];[HouFai53];[PetBar52]. These variations in F0 are typically much smaller in magnitude than those produced by deaf speakers with (vowel-related) erratic pitch; nonetheless, they indicate that some interaction between articulation and phonation does occur, even in the normal situation.

Several mechanisms have been proposed to account for the vowel-related changes in F0 observed in the speech of hearing individuals. Lieberman [Lie70] and Atkinson [Atk73], for example, have argued that the phenomenon is an aerodynamic one, resulting from acoustic coupling between the vocal folds and the supralaryngeal vocal tract. Their arguments are based primarily on computer-simulated, "one-mass" models of the larynx which predict that F0 will be higher when the frequency of the first formant is relatively low, as it is for high vowels such as /i/ and /u/ [FlaLan68];[Fla65]. In these situations, the acoustic input impedance of the vocal tract is typically high in the vicinity of F0, and one would expect that relatively small changes in volume velocity through the glottis would lead to substantial changes in supraglottal (and, thus, transglottal) pressure. According to the coupling hypothesis, these pressure fluctuations across the glottis would serve to increase the frequency of vocal-fold vibration (i.e., to make F0 more like F1) and, as predicted by the computer models, this effect would be greatest when F0 was closest in frequency to F1. (Such acoustic coupling is similar to that which occurs between the resonances of a bugle and a bugler's lips, in that the resonances of the tract effectively dictate to the source the frequencies at which it can vibrate [Ben73].)

An alternative explanation is that vowel related variations in F0 result from changes in vocal-fold tension which occur when the tongue body, the jaw and/or the larynx itself are manipulated during vowel production. Lehiste [Leh70], for example, has noted that muscles of the tongue body and larynx are connected to the superior and inferior parts of the hyoid bone respectively and has suggested that the raised tongue position associated with the production of high vowels might also exert an upward pull on the larynx, thus stretching the vocal folds (longitudinally) and increasing F0. Ohala [Oha77];[Oha73], on the other hand, has argued that tongue pull may lead (via soft-tissue connections) to increased *vertical* tension in the larynx (hence, to higher F0), and Ewan [Ewa75] has noted that this vertical tension might be enhanced by an active lowering of the larynx for the vowel /u/.

Ewan [Ewa79a];Ewa79b] has also suggested that the relatively low F0 associated with low vowels such as /ɑ/ might result, in part, from a "tongue-retraction" or "pharyngeal-constriction" component which would serve to increase the vibrating mass of the vocal folds (by pushing soft tissue down toward the larynx) and hence to decrease F0. Such pharyngeal constriction would also tend to shorten the vocal folds [Gau77];[Lin-Gau72], thereby reducing their tension and contributing to an F0-lowering effect.

Finally, recent electromyographic and X-ray studies [Hon81];[HonBae81] suggest that increases in F0 may be associated with forward movements of the *hyoid bone* which occur when the jaw is stabilized and the tongue root moved forward during the production of high vowels. Honda [Hon81] has argued that such horizontal movements of the hyoid will serve to tilt the thyroid cartilage forward, thus increasing the (longintudinal) tension on the vocal folds and raising F0.

1.3 Deaf vs. Hearing Speakers

Although the relative merits of the mechanisms just described are still being debated (e.g., [ShaPie79]; [OhaEuk78]), each can serve as a useful starting point for examining the exaggerated vowel-related variations in F0 produced by many deaf speakers. Both the acoustic-coupling and the vocal-fold tension hypotheses, for example, assume that small variations in F0 are a natural consequence of *changes* in the configuration of the supralaryngeal vocal tract. Hence, both hypotheses would predict (other things being equal) greater F0 variability among deaf speakers able to produce a relatively wide range of vowels than among those speakers whose articulatory capabilities were more limited. Such findings are essentially the opposite of what would be expected if deaf speakers were using F0 as a *means* of vowel differentiation, as has been suggested by a number of previous investigators [Hor77];[AngKop64].

Whether sufficient acoustic coupling between the larynx and the vocal tract can occur during normal voicing to account for the vowel-related variations in F0 produced by hearing speakers is still open to question [Pet76]. The vocal folds typically vibrate with a complex wavelike motion, opening from bottom to top and closing with considerable force, and it is probable that this type of vibration is relatively insensitive to acoustic changes in the supraglottal system [IshMat68]. Sophisticated laryngeal models, for example, in which each vocal fold is treated as a pair of coupled masses (i.e., "two-mass" models) in order to simulate this wavelike behavior, suggest that F0 is not appreciably influenced by vowel-related changes in the frequency of F1 [IshFla72].

One might speculate, however, that certain inadequate laryngeal postures or modes of vocal-fold vibration (e.g., when the vocal folds are held taut and vibrate only along their margins or when they vibrate without closing completely along their length) might be more susceptible to such acoustic change. In these situations, the introduction of a low frequency supraglottal resonance might sufficiently alter the pressure variations across the larynx (by virtue of a mode of vibration more comparable to that simulated by the original one-mass model [FlaLan68]) to produce a substantial increase in F0. Furthermore, such abnormal modes of voicing might be associated with the "breathy" voice quality often encountered in the speech of the deaf [SteNic79];[Lin76];[Eng62];[Hud37].

There is evidence to suggest that, in certain cases, acoustic coupling will prevent the vocal folds from vibrating at frequencies in the immediate vicinity of F1. A number of experiments have been performed, for example, in which subjects were asked to phonate a vowel into tubes of varying length, thus increasing the effective length of the vocal tract and decreasing F1 [IshFla72];[KagTre37]. The results of these experiments showed that F0 tended to jump sharply upwards as soon as the frequency of the first formant was lowered sufficiently to approach that of the fundamental. This phenomenon was also observed in the computer simulations of the one- and two-mass models referred to above [IshFla72].

Sundberg [Sun75] has suggested that these effects of acoustic coupling on F0 will be more pronounced in cases for which the damping of the vocal tract (in particular, the damping of F1) is relatively low. His argument is motivated, in part, by experiments in which trained singers attempted to intone vowels for which F0 and F1 were approximately equal in frequency. While the two female singers in his study (an alto and a soprano) had little difficulty in accomplishing this task, the male singer (a baritone) was unable to maintain a stable F0 in the vicinity of F1 unless he shifted from chest to falsetto register.

In interpreting these findings, Sundberg drew upon the work of van den Berg [Ber55] who treated the speech-production apparatus as a pair of coupled resonators, one representing the vocal folds (F0) and the other representing the vocal tract (F1). On the basis of this theoretical model, van den Berg argued that the effects of acoustic coupling on F0 control would be negligible in the falsetto register, since the vocal folds typically would not close completely during this type of vibration and, as a result, the damping of the vocal tract would be relatively high. (Acoustical theory predicts that coupling between two such resonators will lead to a type of "repellence" effect which will prevent them from vibrating at exactly the same frequency, the magnitude of this effect being greatest when the frequency distance between the resonances is small and their damping low [JanSun74]; [MorIng68].) Sundberg has used van den Berg's argument, together with the observation that formant bandwidth (hence vocal-tract damping) tends to be greater for female than for male speakers [Fan72];[FujLin71], to account for the dependence of F0 stability (at frequencies near F1) on the register used by and/or the sex of the singers who participated in his experiments.

For most hearing speakers, F0 and F1 tend to be fairly well separated in frequency, and it is unlikely that vocal-fold vibration would be appreciably influenced by such an acoustic-coupling phenomenon [OhaEuk78][PetBar52]. The speech of the deaf, however, is often characterized by an inappropriately high F0 [HorBis72];[Boo66], and it is possible that for certain speech sounds (e.g., high vowels, nasal consonants) and for some deaf speakers, the frequencies of the first formant and the fundamental may be comparable. One might expect, on the basis of the preceding discussion, that this proximity of F0 and F1 could lead to large upward breaks in F0, or to an erratic or unstable F0 contour, if the damping of the vocal tract were sufficiently low. (Hence, in contrast to the situation discussed on page 23, one would expect such deviations in F0 to be less common among deaf speakers with breathy voice quality.)

A number of possible explanations for the vowel-related deviations in F0 produced by deaf speakers can also be advanced on the basis of the "vocal-fold tension" hypotheses outlined in the preceding section. One such explanation (a version of which has been suggested previously by Martony [Mar77];[Mar68]) is that extreme articulatory habits used by some deaf speakers during vowel production may lead to larger-than-normal changes in vocal-fold tension and, thus, to large vowel-related variations in F0. While this hypothesis has yet to be tested with objective data, there is evidence to suggest that some of the articulatory maneuvers thought to influence laryngeal tension in the normal situation may be exaggerated in the speech of many deaf individuals.

Ling [Lin76], for example, has noted that the use of extreme jaw movement tends to be common among deaf speakers and has argued that this movement will also lead to abnormal tongue and/or lip movement during vowel production. He cites the case of a deaf child attempting to produce the high front vowel /i/, starting from a position in which jaw opening is abnormally wide. Among the articulatory adjustments which the deaf child would need to make (in order to produce an "acoustically-acceptable" /i/) is the use of a tongue position which is more fronted (and, perhaps, higher) in relation to the mandible than normal.

There is some indication that such inappropriate jaw positions, together with the resulting distortions of the tongue body, can lead to increased vowel-to-vowel variations in F0. Ohala and Eukel [OhaEuk78], for example, asked hearing speakers to produce a set of test utterances while their jaws were propped open with small wooden bite blocks, and found that the F0 interval between high and low vowels was slightly, but consistently, larger than in the normal situation. The authors attributed this increased F0 variability to an exaggerated pull of the tongue body on the hyoid-larynx complex (and, thus, to an increased laryngeal tension) during the production of high vowels in the bite-block condition.

Cinefluorographic data collected by Boone [Boo66] support the observation that deaf speakers often make use of excessive jaw movement during speech production. Boone's data also indicate that many deaf speakers tend to hold the tongue body in a low and backed position (relative to that assumed by hearing speakers), resulting in a greater-than-normal degree of pharyngeal constriction and, apparently, reducing the range of front-back tongue movement used in vowel articulation [SteNic79]. (The major acoustical consequence of this limited tongue movement is the production of vowels with a relatively narrow range of F2 variability [Mon76], a characteristic of the speech of the deaf which is highly correlated with poor speech intelligibility [Mon78].)

As noted in the preceding section, a tongue-retraction or pharyngeal-constriction component is thought to play a role in lowering F0 during the production of low vowels (e.g., /ɑ/), by reducing the tension on and increasing the vibrating mass of the vocal folds [Ewa79a];[Ewa79b];[Gau77];[Lin-Gau72]. The effect that such an aberrant articulatory posture might have on laryngeal tension (and, thus, F0) during the production of other vowel sounds--in particular, those which would normally require a high and more fronted position of the tongue body--remains to be determined. (Ling [Lin76] has stated that excessive "pharyngeal tension" tends to inhibit F0 variability in the speech of some deaf individuals, by inducing an inappropriate degree of tension in the larynx itself, but this claim is apparently based more on conjecture than on objective data.)

Whether the changes in vocal-fold tension associated with such (exaggerated) articulatory maneuvers or postures would be sufficient to account for the relatively large vowel-related deviations in F0 produced by some deaf speakers is still open to question. However, once again, one might speculate that the deaf speaker's use of an inappropriate laryngeal posture or mode of vocal-fold vibration could serve to exaggerate the influence of small (vowel-related) perturbations in laryngeal tension on F0. In the one-mass model referenced above, for example, F0 was more sensitive to changes in a vocal-cord tension factor when boundary conditions were chosen to simulate hard-walled, elastic as opposed to viscous, inelastic collisions of the vocal folds [FlaLan68].

Finally, although the vowel-related variations in fundamental frequency observed for hearing speakers are relatively small in magnitude, the possibility cannot be ignored that larger changes in F0 are, in fact, a natural consequence of the articulatory maneuvers used in vowel production. Presumably, if this were the case, hearing children would soon learn to anticipate such changes in F0 in their own speech and to compensate for these changes (although only partially) by making appropriate adjustments in laryngeal tension and position [NicSte77]. Deaf children, on the other hand, lacking the necessary auditory feedback, might find such compensation difficult and, thus, continue to produce what would appear to be exaggerated vowel-to-vowel variations in F0.

While the last of these hypotheses would best be tested through a comparison of longitudinal (F0) data for deaf and hearing speakers, such an analysis is beyond the scope of this study. As will be discussed later (Section 4.2.2), however, some information relevant to this hypothesis can be obtained from the literature on the normal development of F0 control.

1.4 Research Objectives and Overview

In recent years, a growing number of researchers and teachers of the deaf has stressed the need for more integrated and theoretically sound procedures for speech diagnosis and training [Boo80];[NicSe79a];[Lin76]. The development of such procedures has been hampered, however, by a poor understanding of the ways in which various problems interrelate in the speech of the deaf and in the mechanisms responsible for these interactions. The literature review presented in Section 1.1, for example, indicates that several contradictory explanations exist for the production of exaggerated vowel-related variations in F0 by deaf speakers. While some investigators have suggested that deaf speakers may use such variations in F0 (rather than articulatory variations) as a *means* of differentiating one vowel sound from another, others have argued that large changes in F0 may be a *consequence* of the deaf speaker's inability to control F0 independently of the articulatory maneuvers used in vowel production.

Although these two explanations clearly have different implications for speech training, neither, as yet, has been well-tested with objective data. One major objective of this thesis, then, is to document more carefully the relationship between the vowel-to-vowel variations in F0 produced by deaf speakers and their proficiency at vowel articulation. A second, and more general, objective of the thesis is to gain a better understanding of the mechanisms responsible for the deaf speaker's problems with vowel-related F0 control through: 1) a systematic collection of acoustic (i.e., F0, formant and spectral) data; and 2) an interpretation of these data in terms of the hypotheses discussed in Section 1.3.

The experimental strategy used in this study was to design a corpus of test utterances in which both vowels and consonants were varied to cover a range of acoustic and articulatory features (e.g., low F1, high tongue position). Groups of deaf and hearing boys and girls were recorded, and the recordings were processed using computer and spectrographic techniques. Details of this experimental procedure, including a description of the acoustic measurements made, are provided in Chapter 2.

Chapters 3 and 4 are devoted to analyzing and interpreting the acoustic data, in line with the research objectives defined above. More specifically, Chapter 3 compares F0 and formant-frequency data for the deaf and hearing boys and girls, and examines these measures relative to intelligibility scores for the deaf speakers. Interactions among vowel articulation, F0 control, severity of hearing loss and a spectral measure indicative of laryngeal posture are also considered in this chapter.

The effects of consonantal context on F0 are examined in Chapter 4, in order to estimate the relative influence of jaw position, tongue height and the frequency of F1 on the deaf speakers' control of F0. The results of these analyses, together with those of the preceding chapter, are then interpreted in terms of the (hypothetical) mechanisms for vowel-related F0 change discussed in Section 1.3.

Chapter 5 summarizes the major findings of the study, and discusses the implications of these findings for the development of speech-training procedures and for future research efforts. Some limitations of the analyses performed in the study are also considered in this chapter.

CHAPTER 2

PROCEDURE

This chapter describes the experimental procedures and data-processing techniques requisite to the analyses carried out in this study. Sections 2.1 through 2.3 provide information about the speakers participating in the study as well as details of corpus design and recording. Section 2.4 describes the methods used in obtaining the acoustic data and intelligibility scores to be examined in later chapters of the thesis.

2.1 Speakers

Twenty deaf and thirteen hearing boys and girls served as speakers. The deaf speakers were all students in the middle and upper schools at The Clarke School for the Deaf in Northampton, Massachusetts, while the hearing speakers were boys and girls living in the greater Boston area. Each of the hearing girls was a volunteer recruited through the High School Studies Program of the Massachusetts Institute of

Technology; the hearing boys were children and acquaintances of faculty members at M.I.T.

Tables 2.1 and 2.2 contain background information on each of the participants in the study. As shown in Table 2.1, the deaf group of speakers was made up of eight boys and twelve girls with profound (83 to 130+ dB) and, in most cases, congenital hearing loss. (As will be seen later, speakers D5 and D6 had relatively poor speech-production capabilities, although each lost his hearing at a fairly late age.) The control group of speakers consisted of five boys and eight girls of comparable age with no known hearing or speech impairment (Table 2.2).

The selection of the deaf speakers taking part in the study was based, to a large extent, on (the author's) subjective judgements of their voice quality and fundamental-frequency control. (These judgements were made after listening to a set of twice-annual recordings of each speaker maintained by the speech department at The Clarke School.) Approximately half of the boys and girls who were chosen were judged to speak with relatively high-pitched (and, occasionally, weak or breathy) voices. The remaining speakers were judged to have stronger, lower-pitched voices and, in the case of some girls, to speak with a "boyish" or "tense" voice quality. An attempt was made to select not only speakers who produced noticeable segmental deviations in F0, but also speakers whose F0 control was more stable (or even monotone).

Table 2.1 - Background Data for the Deaf Speakers

Speaker	Sex	Age (yr:mo)	Hearing Loss* (dB)	Age at Onset** (yr:mo)	Cause***
D1	M	13:1	107	C	U
D2	M	10:10	102	C	R
D3	M	11:5	97	C	H
D4	M	11:10	100	C	R
D5	M	9:7	107	2:6	M
D6	M	10:6	130+	3:9	M
D7	M	9:6	107	C	H
D8	M	10:7	103	0:5	I
D9	F	13:10	83	C	H
D10	F	14:2	107	C	R
D11	F	13:11	107	C	R
D12	F	13:11	112	C	R
D13	F	15:5	93	C	U
D14	F	17:6	100	C	U
D15	F	14:3	100	C	R
D16	F	14:8	83	C	H
D17	F	15:1	98	C	H
D18	F	16:9	90	C	U
D19	F	14:11	110	C	R
D20	F	12:2	100	C	K

* Standard pure-tone hearing threshold measurements averaged for 500, 1000 and 2000 Hz for the better ear.

** C = congenital

*** H = hereditary; K = Rh or jaundice; M = meningitis;
R = prenatal rubella; I = other infections;
U = unknown

Table 2.2 - Background Data for the Hearing Speakers

Speaker	Sex	Age (yr:mo)
H1	M	10:6
H2	M	10:1
H3	M	13:6
H4	M	10:2
H5	M	10:4
H6	F	15:2
H7	F	16:10
H8	F	15:3
H9	F	15:11
H10	F	13:0
H11	F	15:10
H12	F	12:8
H13	F	16:1

In order to avoid the confounding influence of adolescent voice breaks (associated with rapid laryngeal growth) on the fundamental-frequency data, all of the boys chosen for the study were prepubertal. As shown in Tables 2.1 and 2.2, the deaf boys ranged in age from 9 years, 6 months to 13 years 1 month and the hearing boys from 10 years, 2 months to 13 years, 6 months at the time of recording. The deaf and hearing girls, on the other hand, were somewhat older--ranging in age from 12 years 2 months to 17 years, 6 months and from 12 years, 8 months to 16 years, 10 months respectively. Since laryngeal growth during adolescence is much less pronounced for girls than for boys [Kah75], age-related breaks in voicing were not expected to be a problem for the female speakers.

2.2 Corpus

The corpus of utterances analyzed in this study consists of forty monosyllabic nouns in which both vowels and consonants are systematically varied. During the recording sessions, these nouns were spoken as the second word of a two-word phrase (i.e., in the context "the ____"). As shown in Table 2.3, each of the nouns in the test corpus is of the form C1-V-C2 (except box) where V is one of nine English vowels (/i,I,ɛ,æ,ɑ,ʌ,ɝ,ʊ,u/) or the diphthong /ɑy/. The target nouns can also be divided into four categories depending on the manner, place and voicing characteristics of

Table 2.3 - Test Corpus

Monosyllabic Nouns

Vowel Target	Category 1	Category 2	Category 3	Category 4
	[bVC2]*	[C1VC2]**	[C1VC2]***	[C1VC2]****
/i/	beak	peak	----	leak
/I/	bid	pit	king mitt	lid
/ɛ/	bed	pet	net	well
/æ/	bag	cat	man	rag
/ɑ/	box	top	dawn	wall
/ʌ/	bug	cut	mug	rug
/ɝ/	bird	dirt	burn	word
/ʊ/	book	cook	nook	wool
/u/	boot	toot	moon	root
/ɑy/	bike	pipe	time	light

* C2 = voiced stop (5); unvoiced stop (4) (exception: box)

** C1,C2 = unvoiced stop (exception: dirt)

*** C1 = nasal; C2 = stop (4)
C1 = stop; C2 = nasal (4)
C1,C2 = nasal (2)

**** C1 = /r/ (3); /l/ (3); /w/ (4)
C2 = /l/ (3); voiced stop (4); unvoiced stop (3)

Note: Numbers in parentheses indicate the number of occurrences. Exceptions (i.e., king, box, dirt) were necessary in order to form meaningful nouns (see text).

C1 and C2. Details of this division are provided in the lower section of Table 2.3; however, it can be noted briefly here that the nouns of Categories 1 and 2 contain only stop consonants while those of Category 3 contain at least one of the nasal consonants /m,n,ŋ/ and those of Category 4 at least one of the sonorants /r,l,w/.

Two objectives were pursued in selecting the consonantal phonemes included in the test corpus. First, an attempt was made to choose consonants characterized by a range of articulatory and acoustic features, including features relevant to the production of high vowels. As will be discussed in Chapter 4, these consonants were used in examining the relative influence of high jaw position (labial stops), high tongue position (velar stops) and a low-frequency first formant (nasals and sonorants) on the deaf speakers' control of F0.

The second objective in designing the test corpus was to select consonants which, when combined with the various vowel targets, formed nouns which were familiar to even the youngest deaf speakers. The choice of such nouns was motivated by the assumption that a more natural manner of speaking would be encouraged by the use of meaningful test utterances.

2.3 Recording Techniques

Each of the speakers in the study was recorded using a voice microphone and two miniature accelerometers [SteKal75]. The microphone was suspended at a fixed distance from the speaker's mouth, and the two accelerometers were attached by double-stick tape to the speaker's nose and throat respectively. (All three of the recording devices were connected to a headband which the speaker adjusted comfortably around his/her head.) In this manner, three signals were simultaneously recorded: 1) the speaker's voice; 2) the amount of vibration at the side of the speaker's nose; and 3) an estimate of the speaker's glottal output during voiced segments of the recordings. A second microphone was also used to record the experimenter's comments and labelling information.

All of the recordings were made in sound-treated rooms, either at The Clarke School for the Deaf or in the Speech Communications Laboratory at the Massachusetts Institute of Technology. The overall experimental set-up (as described above) was comparable in both locations; the specific pieces of equipment used were as follows:

<u>Clarke School:</u>

accelerometers:	Bolt, Beranek and Newman Model 501
voice microphones:	Thermo Electron electret condenser microphone, Model 5333

tape recorder: SONY four-channel recorder, Model TC-854-4

M.I.T.

accelerometers: Bolt, Beranek and Newman Model 501
voice microphones: Thermo Electron electret condenser microphone, Model 5333 (speakers); Shure Unidyne III, Model 545 cardiod microphone (experimenter)
tape recorder: TEAC four-channel recorder, Model 3340

The TEAC recorder was also used in conjunction with a Braun stereo taperecorder (Model TG 1000) in copying the original four-channel recordings to a set of two-track tapes suitable for computer analysis (Section 2.4).

Appropriate adjustments to the equipment were made at the start of each recording session (e.g., record levels were set and accelerometer placement was adjusted), while the speaker read a set of practice utterances similar to those contained in the test corpus. The test phrases were then presented to the speaker one by one (on index cards), at a rate determined by the experimenter. Labelling information was inserted between phrases, helping to avoid a "list" effect in the speakers' intonation contours.

If a speaker stumbled over or consciously mispronounced one of the test utterances during a recording session, he or she was asked to repeat the phrase correctly. However, in order to encourage as natural and relaxed a manner of speaking as possible, no attempt was made to correct phonemic errors produced unwittingly by the deaf speakers. This latter precaution was taken not only to avoid the confounding influence of increased emotional tension on the F0 data, but also because examining certain limitations in the deaf speakers' articulatory skills (e.g., vowel neutralization) was an important part of the present study.

2.4 Data Processing

2.4.1 Fundamental Frequency

Fundamental-frequency contours were obtained for each of the test phrases using the FPRD and F0PLOT programs developed by Henke [Hen76]. These programs were implemented on the PDP-9 computer facility of the Speech Communications Laboratory at M.I.T. A block diagram of this facility, which consists of the PDP-9 computer and a set of peripheral devices appropriate for speech analysis, is shown in Figure 2.1.

The FPRD program uses a zero-crossing algorithm, together with a set of user-defined tracking parameters, to estimate the fundamental period (T0) of a digitized speech waveform. Fundamental frequency is then calculated by FPRD (F0 = 1/T0 x

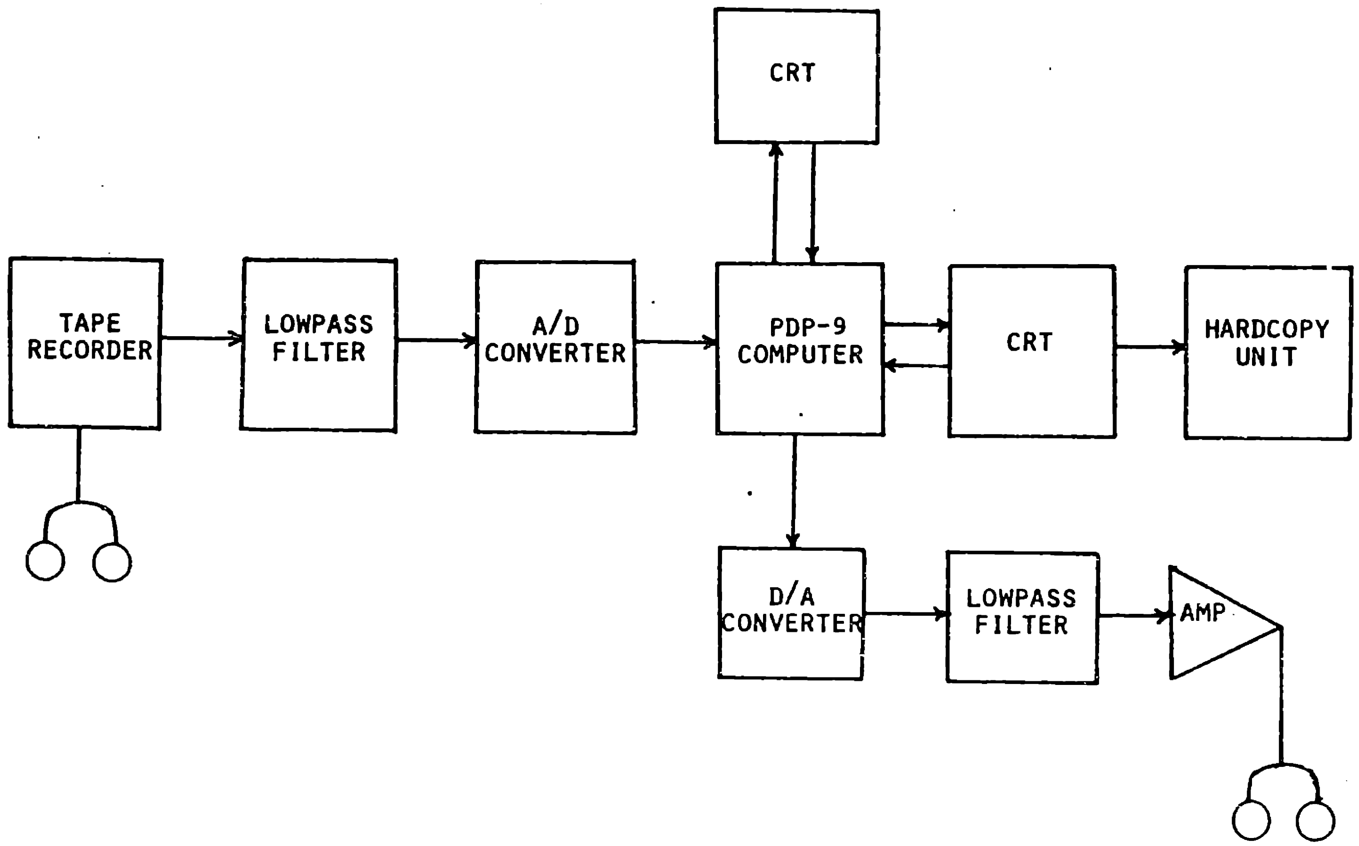


Figure 2.1 - Block diagram of the computer facility.

sampling rate) and displayed automatically as a function of time (on one of the CRT screens) by the F0PLOT program. A display of the relative amplitude of the digitized waveform as a function of time is also provided.

Several features of the FPRD and F0PLOT programs made them well-suited to the analyses of the present study. First, the tracking parameters mentioned above are set by the user to accommodate the range of fundamental frequencies and the mean F0 most likely to be used by any given speaker. Thus, with the proper choice of parameters, the FPRD program can be made to perform well for both high and low-pitched voices and for speakers using either small or large variations in F0.

The FPRD program also assigns a confidence rating to each apparent period boundary in the speech waveform, ranging from 1 (not at all confident) to 4 (very confident that a period boundary is present). When the confidence ratings are high, the contour plotted by the F0PLOT program is an unbroken (and relatively smooth) curve during voiced segments of the speech signal. When the ratings are low, the contour appears as a set of disconnected points, the presence of which indicates either a poor choice of tracking parameters or some aperiodicity in the speech waveform. The user can decide between these latter two sources of error by examining a replica of the processed waveform, which is saved by the FPRD program and displayed on one of the CRT screens. This capability was especially useful in the analysis of utterances characterized by erratic breaks in voicing or by sharp changes in F0, as was sometimes necessary for the deaf speakers.

In the analyses of the present study, F0 contours were obtained by processing the throat accelerometer signal, since this waveform provided a somewhat cleaner estimate of glottal output than that recorded by the voice microphone (i.e., it was less influenced by the frequency of F1). The voice microphone signal was used, however, in monitoring the phrases which were input to the computer. (As noted earlier, two-track tapes suitable to this type of analysis had been made from the original four-channel recordings.)

The procedures used in obtaining the F0 data for any given phrase were as follows (see also the block diagram of Figure 2.1). The throat accelerometer signal was lowpass filtered at approximately 900Hz, sampled at 10 KHz and converted from analog to digital form. The FPRD and F0PLOT programs were then used to compute and display an F0 contour for the utterance, and the contour was checked for errors of the types described above. If necessary, the tracking parameters were reset and the signal processing repeated. When an acceptable F0 contour had been obtained, a hardcopy of the display was made and all subsequent measures of F0 were taken from this hardcopy.

Quantitative measurements of F0 were made (to the nearest 5 Hz) at two or more locations in the target syllables, depending upon consonantal context. These locations included the following:

1) the "center" of each vowel target
2) the "center" of each nasal and sonorant (i.e., /m,n,ŋ,r,l,w/)
3) points approximately 10 and 20 milliseconds following vowel "onset" in syllables beginning with voiced and voiceless stop consonants (i.e., /b,d,g,p,t,k/).

Points on the hardcopy F0 contours corresponding to these locations were determined in the manner described below.

For most nouns beginning with stop consonants, a break in voicing occurred during consonant closure, and the portion of the F0 contour corresponding to the target syllable appeared as a discrete segment on the hardcopy display. (As noted earlier, each of the target nouns was recorded in the carrier phrase "the ____".) In these situations, the "onset" of the target vowel was defined as the first connected point on the appropriate segment of the F0 contour, and the offset of the vowel was defined in a similar manner (i.e., as the last connected point on the F0 contour) if the phoneme following the vowel was a stop consonant. (F0 at such points was computed by FPRD using the first and last sets of period boundaries in the vowel waveform to which a confidence rating of 4 had been assigned.) The point on the F0 contour lying half way between vowel onset and offset was then defined as the "center" of the vowel target.

In the case of syllables containing nasals or sonorants (or syllable-initial stop consonants for which voicing continued throughout closure), the onsets and offsets of the various speech sounds were determined using wideband spectrograms and segmentation rules similar to those of Peterson and Lehiste [PetLeh60]. (The spectrograms were made at M.I.T. using a Voiceprint Laboratories sound spectrograph, Model 4691A.) These speech-sound boundaries were located on the corresponding F0 contours, and the "center" of each vowel and consonant was defined as above (i.e., as the point on the contour halfway between the onset and offset of the appropriate segment).

Measures indicative of the mean F0 and the amount of vowel-related variability in F0 used by the deaf and hearing speakers were computed on the basis of F0 measurements made at the first of the three locations defined above (i.e., at the center of the vowel targets). F0 data from all three locations, together with qualitative descriptions of the shape and stability of the F0 contours, were used in examining the influence of consonantal context on the deaf speakers' control of F0. Details of these analyses will be provided in the appropriate sections of Chapters 3 and 4.

2.4.2 Formant Tracking and Short-Term Spectra

Formant-frequency data and short-term spectra were obtained using a linear-predictive-coding formant tracker (LINPC) developed by Goldstein [Gol79];[Gol76] and implemented on the PDP-9 computer facility described earlier. The LINPC program uses a pitch-asynchronous, covariance-type analysis, to perform a frame-by-frame calculation of predictor coefficients for a selected segment of speech waveform. Pole locations of the all-pole transfer function are then computed for each frame (by means of a root-finding algorithm) and converted to a set of formant frequencies and bandwidths. The formant tracks are plotted automatically as a function of time on one CRT screen (see Figure 2.1), and numerical values of both formant frequencies and bandwidths can be viewed on the other. An option in the program also allows for the computation and display of LPC and DFT (Discrete-Fourier-Transform) spectra for any frame of interest.

A number of parameters in the LINPC program can be set by the user, depending upon characteristics of the speech being processed and upon the time and frequency resolution required. These parameters include: 1) the number of predictor coefficients calculated; 2) the maximum bandwidth of allowable formants; 3) the number of sample points per frame; and 4) the step size (in samples) between adjacent frames. In the analyses of the present study, the latter two parameters were held constant at 200 and 100 samples respectively. Thus, with a sample frequency of 10 KHz, each frame analyzed corresponded to a 20-millisecond segment of the speech waveform, and adjacent frames overlapped by 10 milliseconds.

The formant-bandwidth criterion was also held constant at the program's default value of 700 Hz. In most instances, this value proved adequate for including all important spectral peaks in the formant array. (The performance of the LINPC program was checked for each vowel processed by comparing the computed formant locations with LPC and DFT spectra for relevant frames, as defined below.) Occasionally, however a formant appeared to be incorrectly eliminated, and information about its frequency and bandwidth had be recovered from a temporary storage location maintained by the LINPC program.

Unlike the other parameters, the number of predictor coefficients calculated was varied both across speakers and across utterances for a given speaker, depending upon age, sex and voice quality. (Again, the criterion used in determining the number of coefficients calculated was a good match between computed pole locations and the vowel spectra.) For the most part, 10 or 12 predictors were found to be adequate for non-nasalized, non-breathy vowels; the number of predictors needed was relatively higher (12 to 16), however, when nasal (or suglottal) poles and zeros appeared in the vowel spectra or when the amplitude of the first harmonic was excessively high.

The specific procedures used in obtaining the formant data for a given vowel target were the following (see also Figure 2.1). The voice microphone signal was lowpass filtered at 5 KHz, sampled at 10 KHz and preemphasized digitally at 6 dB/octave. The digitized speech waveform was viewed on one of the CRT screens, and a 100-millisecond segment centered approximately at the point in the vowel at which the F0 measurement had been made (i.e., the vowel "center" as defined earlier) was marked for analysis. The LINPC program was then used to compute and display formant locations for this segment of the vowel target, and frequency values corresponding to F1 and F2 were averaged over the three frames nearest the center of the segment. These averaged estimates of F1 and F2 (rounded off to the nearest 5 Hz) are the data upon which the relevant analyses of Chapter 3 are based.

LPC and DFT spectra (for one or more of the three center frames) were used in judging the adequacy of the computed formant values, in the manner described above. These spectra were also used, together with phonetic transcriptions of the target utterances (Section 2.4.3), in deciding which poles to choose as formants when extraneous peaks appeared in the vowel spectra. The most common source of extraneous spectral peaks was vowel nasalization. Occasionally, however, the amplitude of the first harmonic was excessively high (for non-nasalized vowels), and the program mistakenly identified this harmonic as F1, F1 as F2 and so on.

In a few cases, the LINPC program identified two poles in the vicinity of F1 when only one peak appeared in the vowel spectrum, despite adjustments in the number of coefficients calculated. (This problem may have been due to a poor choice of the formant-bandwidth criterion.) For some of these vowel tokens, the "correct" location of F1 appeared to be midway between these two poles, and their frequencies were averaged to obtain an approximate value of F1. Tokens for which a reasonable estimate of F1 or F2 could not be made were discarded.

Finally, hardcopies were made of spectra computed at the centermost frame of each vowel target. These hardcopies were used in estimating the relative amount of low-frequency energy in the vowel spectra and in examining the amount of spectral noise at higher harmonics. Both measures were used in inferring the laryngeal postures maintained by the deaf speakers, in a manner to be discussed in Chapter 3.

2.4.3 Other Relevant Data

Intelligibility scores and phonetic transcriptions of the target utterances were also obtained for the deaf speakers. The phonetic transcriptions were made at M.I.T. by three trained linguists, two of whom transcribed the utterances of each deaf speaker. These transcriptions were used both in identifying the probable locations of F1 and F2 when the vowel spectra were ambiguous (see Section 2.3.2) and in differentiating between correct and incorrect productions of consonantal phonemes. As will be discussed in Chapter 4, this latter type of information was important in examining the influence of different articulatory configurations (e.g., different places of consonant articulation) on the deaf speakers' control of F0.

The intelligibility scores, which are listed in Table 2.4, are based on annual measures of speech intelligibility made for all students at The Clarke School. Each of these annual measures represents the percentage of syllables, in a set of six Magner sentences [Mag72], that was correctly identified by a panel of (six or more) naive listeners. The scores listed in Table 2.4 were computed by averaging the percentages obtained for each speaker for the three years prior to recording. These intelligibility data are used in the present study as an independent, subjective measure of the deaf speakers' articulatory skills, as discussed more fully in Chapter 3.

Table 2.4 - Intelligibility Data for the Deaf Speakers

Speaker	Intelligibility Score (%)
D1	73.0
D2	91.3
D3	92.0
D4	71.7
D5	22.0
D6	60.0
D7	23.3
D8	28.3
D9	59.7
D10	76.3
D11	44.0
D12	29.0
D13	97.0
D14	41.7
D15	65.3
D16	95.0
D17	60.0
D18	86.3
D19	27.7
D20	32.3

CHAPTER 3

ARTICULATORY PROFICIENCY AND LARYNGEAL CONTROL

This chapter examines the relationship between the vowel-to-vowel variations in F0 produced by the deaf speakers and their proficiency at vowel articulation. This is accomplished through a comparison of F0 and formant-frequency data for the deaf and hearing boys and girls and through an examination of these data relative to intelligibility scores for the deaf speakers. Interactions among vowel articulation, F0 control, severity of hearing loss and a spectral measure indicative of laryngeal posture are also considered in this chapter.

3.1 Vowel Articulation and F0 Control

3.1.1 Defining the Relevant Measures

Before the data are presented, it will be useful to define a number of measures which are discussed in this section. These measures are based on the F0 and formant-frequency measurements made at the "centers" of the vowel targets (Section 2.4.1) and provide estimates of the mean F0, the relative amount of F0 variability and the range of F1 and F2 values used by each of the deaf and hearing speakers. The measures include the following:

1) "mean F0" = F0 averaged over the forty vowel targets

2) " ΔF0" = cross-vowel range of $\frac{F0(V)}{F0(/\text{ɑ}/)}$

$$\frac{F0(V)}{F0(/\text{ɑ}/)} = \frac{\text{mean F0 for a given vowel target}}{\text{mean F0 for the vowel /ɑ/}}$$

3) "range F1" = cross-vowel range of F1(V)

F1(V) = mean F1 for a given vowel target

4) "range F2" = cross-vowel range of F2(V)

F2(V) = mean F2 for a given vowel target

Measures 1 through 4 are listed for each of the deaf and hearing speakers in Table 3.1 while values of F0(V), F0(V)/F0(/ɑ/), F1(V) and F2(V) as a function of the vowel target (V) are provided in Appendix A. It should be noted that a normalized measure of F0 (i.e., F0(V)/F0(/ɑ/)) is used in computing the relative amount of F0 variability used by

Table 3.1 - F0 and Formant-Frequency Data for Individual Deaf and Hearing Speakers

speaker	ΔF0	mean F0 (Hz)	range F1 (Hz)	range F2 (Hz)
D1	.403	221	653	702
D2	.359	251	598	909
D3	.265	314	702	1640
D4	.175	300	599	838
D5	.135	280	370	433
D6	.132	270	153	383
D7	.126	321	548	1025
D8	.092	360	402	695
D9	.562	282	398	1180
D10	.520	270	434	667
D11	.500	221	246	656
D12	.490	250	324	518
D13	.347	239	526	1492
D14	.294	255	628	1122
D15	.292	288	401	660
D16	.211	221	384	1383
D17	.210	298	256	417
D18	.161	206	491	1217
D19	.139	308	534	389
D20	.123	345	294	879
range	.562-.092	206-360	153-702	383-1640
H1	.195	234	454	1377
H2	.165	220	390	1561
H3	.126	189	436	1827
H4	.116	222	267	1740
H5	.061	223	501	1645
H6	.139	187	600	1378
H7	.127	232	428	1364
H8	.112	207	400	1253
H9	.081	214	301	1136
H10	.077	207	513	1319
H11	.071	197	320	1099
H12	.062	222	364	1321
H13	.037	223	265	1188
range	.037-.195	187-234	265-600	1099-1827

individual speakers, in order to compensate for differences in the speakers' overall F0 levels [Pet76]. This measure is also used in comparing the magnitude and direction of vowel-to-vowel variations in F0 produced by groups of deaf and hearing speakers, as will be discussed in the following sections.

3.1.2 Averaged Data vs. Data for Individual Speakers

Table 3.2 compares averaged values of F0, F1 and F2 as a function of the vowel target for the deaf and hearing groups of speakers. (Again, these data are based on the F0 and formant measurements made at the "center" of each vowel target, as defined in Secton 2.4.1.) As a group, the deaf speakers produce a wider range of cross-vowel fundamental frequencies than do the hearing speakers (48 Hz vs. 15 Hz), a comparable range of first-formant frequencies (405 Hz vs. 388 Hz) and a much smaller range of second-formant frequencies (663 Hz vs. 1373 Hz). F0 is also higher for all vowels for the deaf group of speakers.

These averaged data are, for the most part, similar to those reported by Angelocci, Kopp and Holbrook [AngKop64] and by Horwich [Hor77]. (One exception is that the range of F1 values produced by the deaf speakers in the latter studies was considerably less than normal.) Several observations suggest,

Table 3.2 - Averaged F0 and Formant-Frequency Data

(All measurements are in Hz.)

vowel target	deaf speakers (n=20)			hearing speakers (n=13)		
	F0	F1	F2	F0	F1	F2
/i/	302	510	2153	222	387	2612
/I/	286	525	2094	218	496	2218
/ε/	266	704	1889	211	676	2017
/æ/	263	784	1884	211	775	2060
/ɑ/	254	833	1511	213	738	1239
/ʌ/	263	790	1585	207	702	1515
/ɝ/	269	572	1716	210	534	1600
/ʊ/	278	597	1513	218	547	1262
/u/	300	428	1490	219	410	1411
/ɑy/	270	729	1867	208	688	1814
range	48	405	663	15	388	1373

however, that these authors' hypotheses about a trade-off between F0 variability and articulatory skill may be inappropriate for many of the deaf boys and girls.

First, examination of the data for individual speakers in Table 3.1 shows that large differences exist among the deaf boys and girls with respect to cross-vowel variability in F0 and to the range of F1 and F2 values produced. While some of the deaf speakers produce a relatively wide range of F0, F1 *and* F2 values (e.g., speakers D3 and D13), others produce little variation in F0, F1 *or* F2 (e.g., speakers D8 and D19). Differences in mean F0 are also large among the deaf speakers.

Second, although the amount of F0 variability used by the deaf speakers is greater, on average, than that used by the hearing controls, the *direction* in which F0 varies as a function of vowel height is comparable for the two groups of speakers. This is illustrated graphically in Figure 3.1, in which F0 (normalized relative to each speaker's mean F0 for the vowel /ɑ/) is plotted as a function of vowel target. The similarity in the shapes of the curves for the deaf and hearing speakers suggest that the vowel-related variations in F0 produced by the two groups may be *qualitatively* similar and, presumably, related to the articulatory maneuvers used in vowel production. This (presumed) relationship between F0 variability and vowel articulation will be examined in the following section.

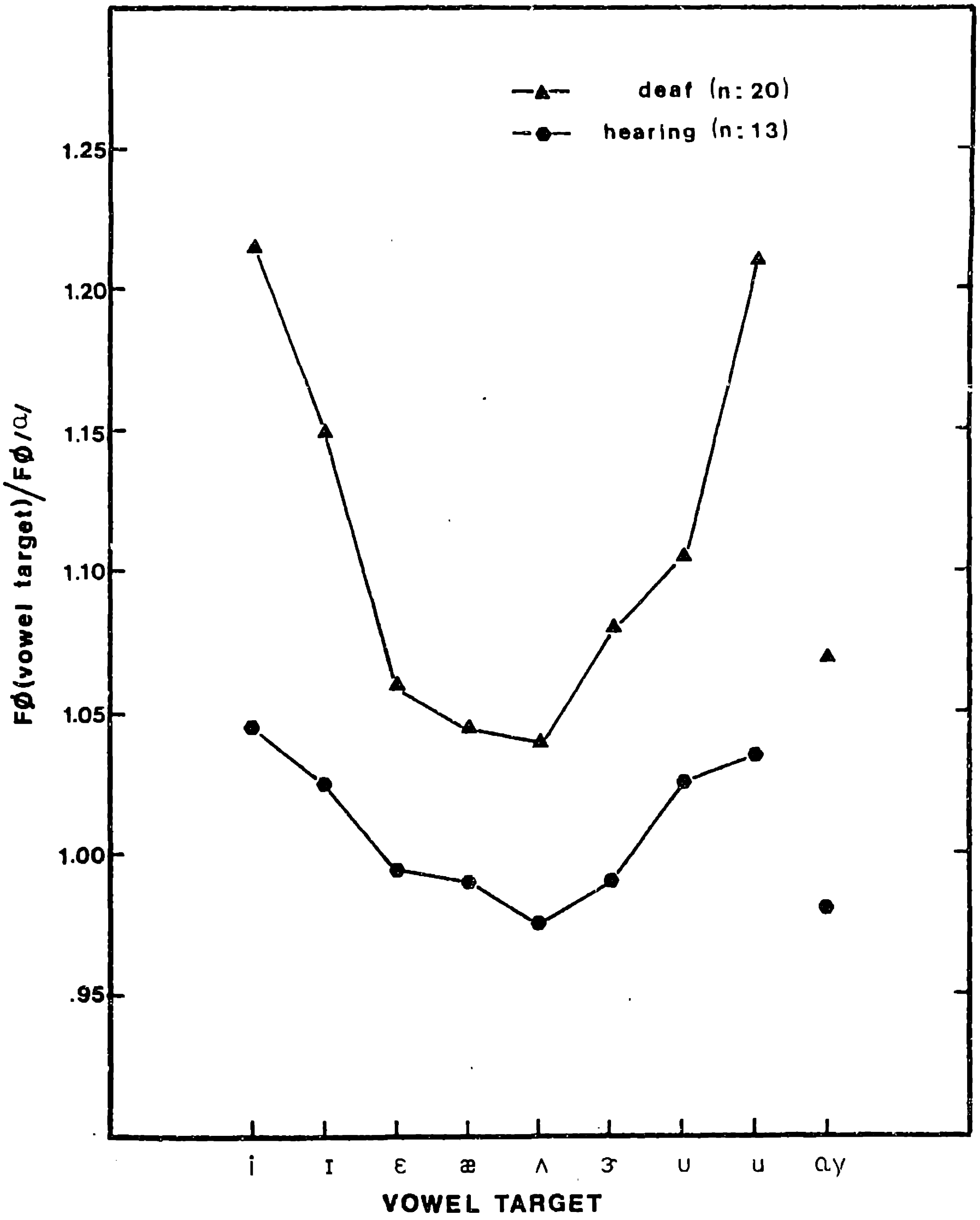


Figure 3.1 - Normalized FØ as a function of the vowel target for the deaf and hearing speakers.

3.1.3 Grouped Data

In order to examine the relationship between F0 variability and articulatory proficiency more carefully, a criterion for grouping the deaf speakers is needed. Simply dividing the deaf speakers into groups on the basis of the amount of vowel-related F0 variability which they produce does not prove very useful, as a one-to-one correspondence between F0 variability and and formant variability does not exist. Compare, for example, the data for speakers D16 and D17 in Table 3.1. While the amount of vowel-to-vowel variability in F0 (i.e.,"△F0") used by the two speakers is approximately the same, the range of F1 values produced by speaker D16 is 1.5 times as large and the range of F2 values 3.3 times as large as those produced by speaker D17.

A criterion which takes into account both F0 variability and mean F0 proved to be more useful in categorizing the deaf speakers. As shown in Figure 3.2, the deaf boys and girls can be divided into three groups (A,B,C) when mean F0 is plotted as a function of vowel-to-vowel variability in F0 (" △F0"). (The rationale for this particular grouping will become apparent in the discussion which follows.) The remainder of this section is devoted to describing the differences in the F0 data for these three groups of speakers and to examining the relationship between F0 variability and articulatory skill

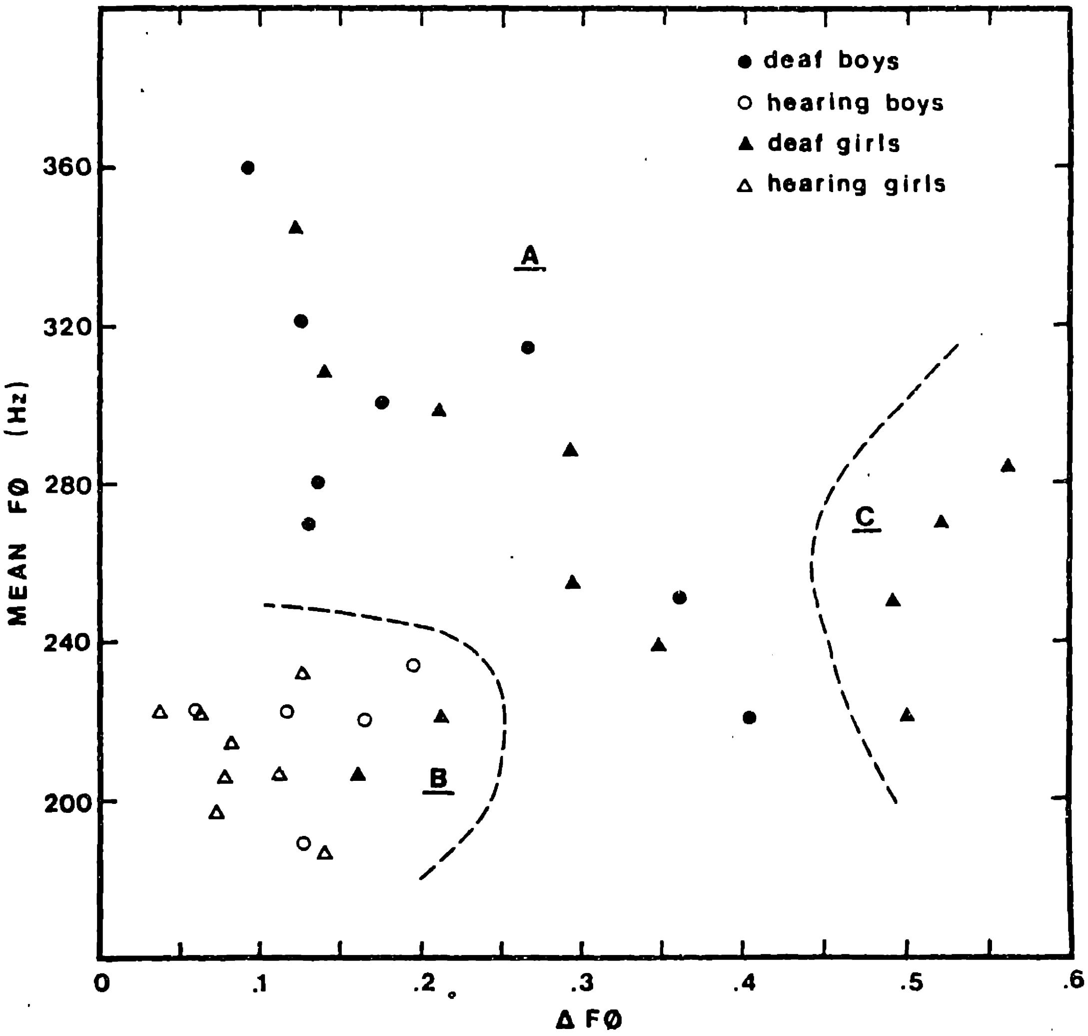


Figure 3.2 - Scatter plot of mean F0 versus vowel-to-vowel variability in F0 for the deaf and hearing boys and girls. (See text for an explanation of Groups A, B and C.)

for the speakers within each group.

Group A:

(Speakers D1-D4, D13-D15, D17, D19-D20)

The majority of deaf speakers, eight boys and six girls, fall into Group A. Among these speakers, a high inverse correlation (r=-.797) exists between mean F0 and F0 variability ("△F0"). Examination of Figure 3.2 indicates that those speakers for whom F0 variability is comparable in magnitude to normal (i.e., to that produced by the hearing speakers) tend to use a mean F0 that is inappropriately high, while those speakers for whom mean F0 is closer to normal tend to produce greater-than-normal vowel-to-vowel variations in F0.

This relationship can be observed more clearly in Figures 3.3 and 3.4, in which the fourteen deaf speakers in Group A have been divided into two subgroups on the basis of age and sex: Group A-1 consists of the the three girls and four boys with the lowest "△F0" scores and Group A-2 consists of the three girls and four boys with the highest "△F0" scores. (This somewhat arbitrary criterion for subdividing Group A was chosen in order to allow separate comparison of formant-frequency data for boys and girls, as discussed below.) Figure 3.3 represents a plot of normalized F0 as a function of vowel target for Groups A-1, A-2 and for the hearing controls. (As in Figure 3.1, the F0 data were normalized relative to each speaker's mean F0 for the vowel /ɑ/ and then averaged.) Figure 3.4 presents non-normalized F0 data for the same three groups of speakers.

Comparison of these two figures shows that, although the deaf speakers of Group A-1 use an amount of vowel-related F0 variability that is comparable to (or slightly less than) that used by the hearing speakers (Figure 3.3), they tend to produce all vowels with much higher F0 than normal (Figure 3.4). The speakers of Group A-2, on the other hand, use a much larger amount of F0 variability and, in particular, tend to produce the vowels /i,I,u/ (and possibly /U/) with excessively high F0 relative to /ɑ/ (Figure 3.3). While F0 tends to be somewhat higher than normal for all vowels for the speakers of Group A-2, it is never as high as that for the speakers of Group A-1 (Figure 3.4).

Figures 3.5 and 3.6 compare the "articulatory proficiencies" of the speakers in Groups A-1 and A-2 (in terms of F1 and F2 measurements made at the center of each vowel target except /ɑy/) with those of the hearing boys and girls. In these figures, vowel diagrams for the male and female speakers are plotted separately. This division was made in order to account for the influence of anatomical differences (due to differences in both the age and sex of the speakers) on the formant data and, possibly, on the mechanisms

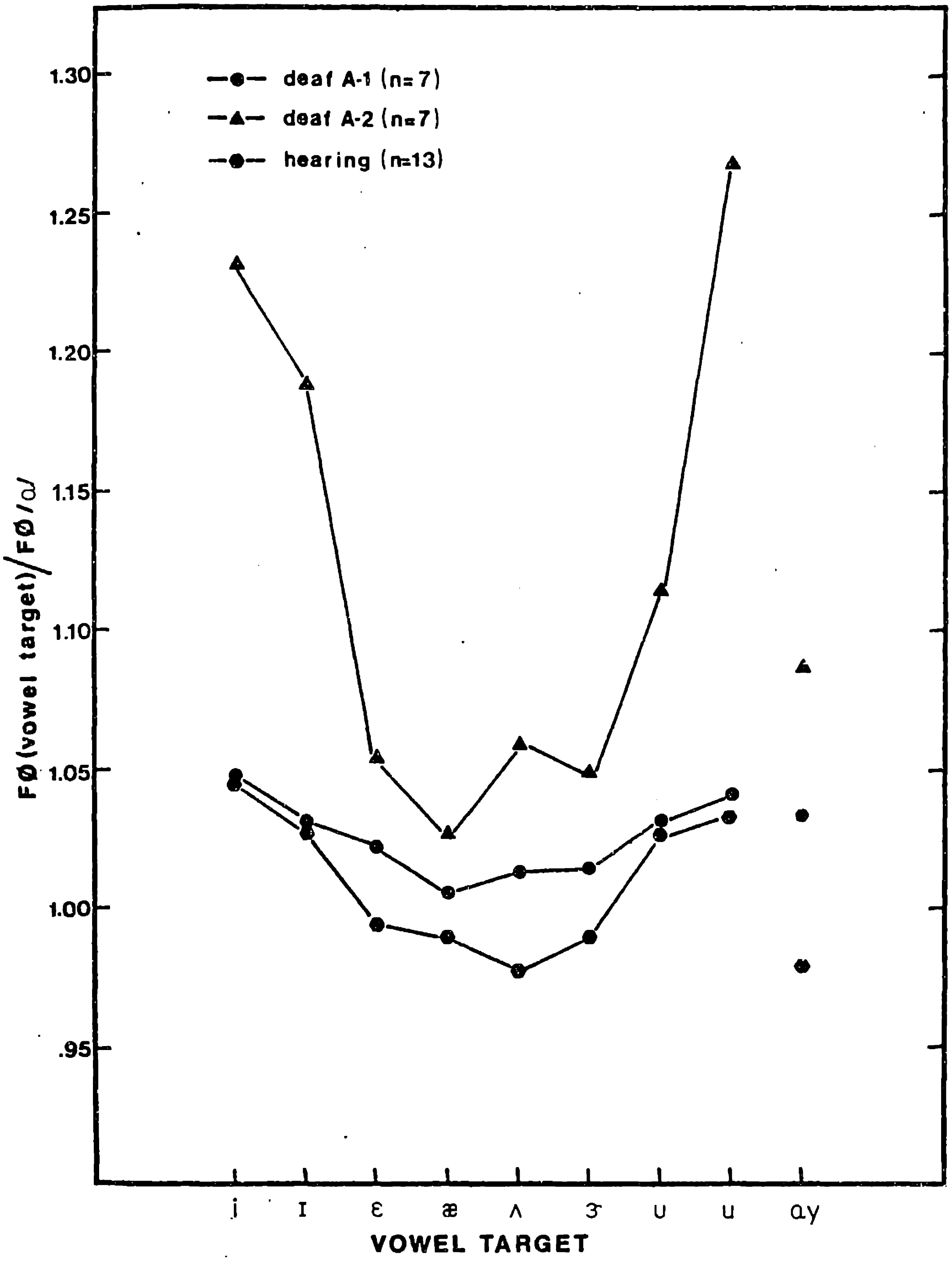


Figure 3.3 - Normalized FØ as a function of the vowel target for the deaf speakers in Groups A-1 and A-2 and for the hearing controls.

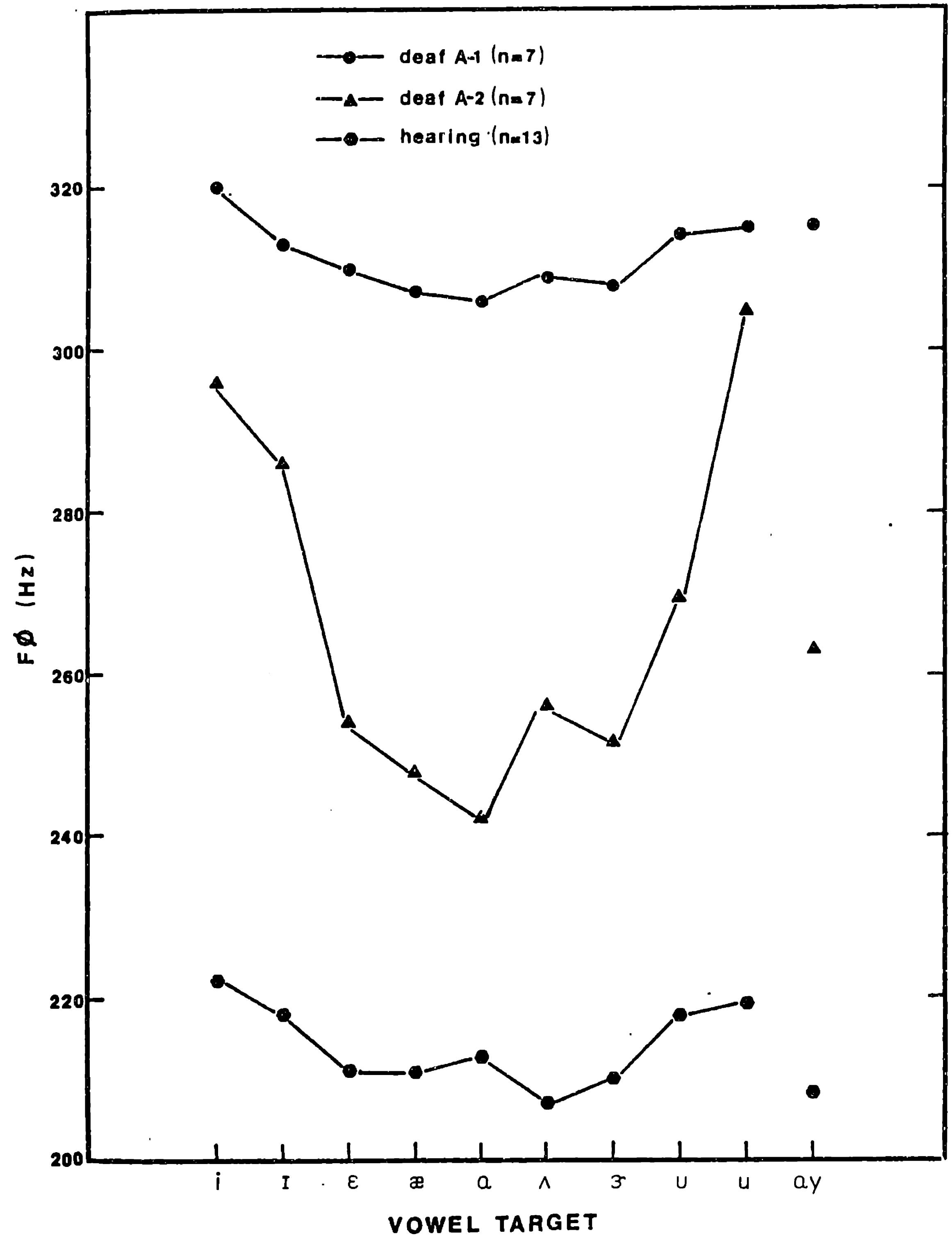


Figure 3.4 - Non-normalized FØ as a function of the vowel target for the deaf speakers in Groups A-1 and A-2 and for the hearing controls.

responsible for interactions between vowel articulation and F0. It should be noted that the vowel diagrams for the hearing boys and girls are similar in shape, but that the girls tend to produce a somewhat smaller range of F1 and F2 values.

The formant-frequency data shown in Figure 3.5 indicate that the range of F1 and F2 values used by the deaf speakers of Group A-1 (i.e., those speakers who produce relatively small cross-vowel variations in F0) is substantially reduced relative to normal. In particular, F1 is never as low in frequency (at least on average) as it is for the hearing boys and girls, suggesting that this group of deaf speakers may be incapable of (or may avoid) producing vowels with the high tongue position appropriate to the high vowels /i/ and /u/.

On the other hand, the vowel diagrams of Figure 3.6 show that the deaf speakers of Group A-2 (i.e., those who produce relatively large vowel-to-vowel variations in F0) tend to produce a much wider range of F1 and F2 values than those in Group A-1. While the range of F2 values used by this group of speakers (i.e., Group A-2) is slightly reduced relative to normal, the range of F1 values (especially for the deaf boys) tends to be larger than that used by the hearing speakers.

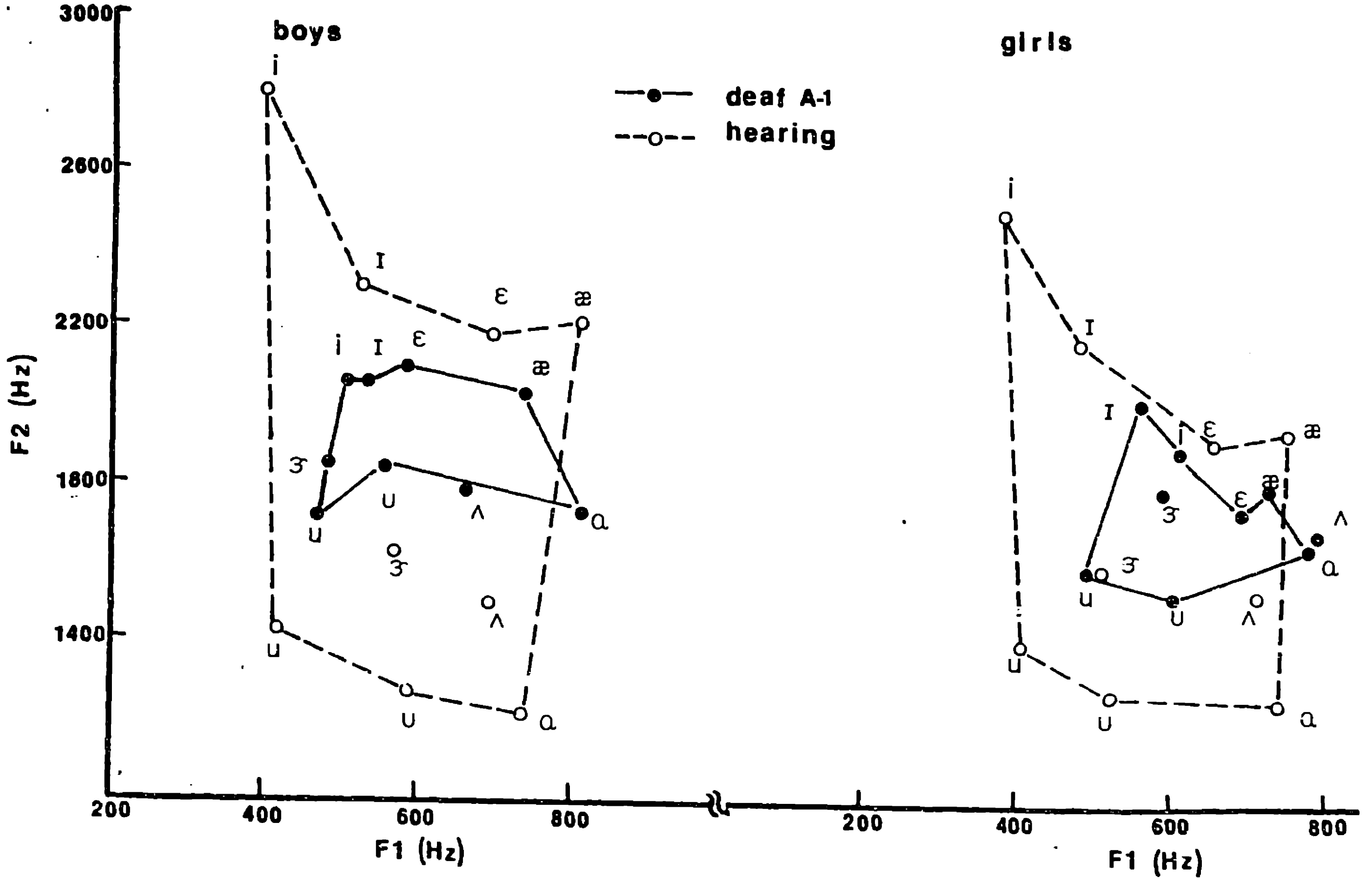


Figure 3.5 - Vowel diagrams (F2 versus F1) for the deaf speakers of Group A-1 and for the hearing controls. Formant data are plotted separately for boys (left) and for girls (right).

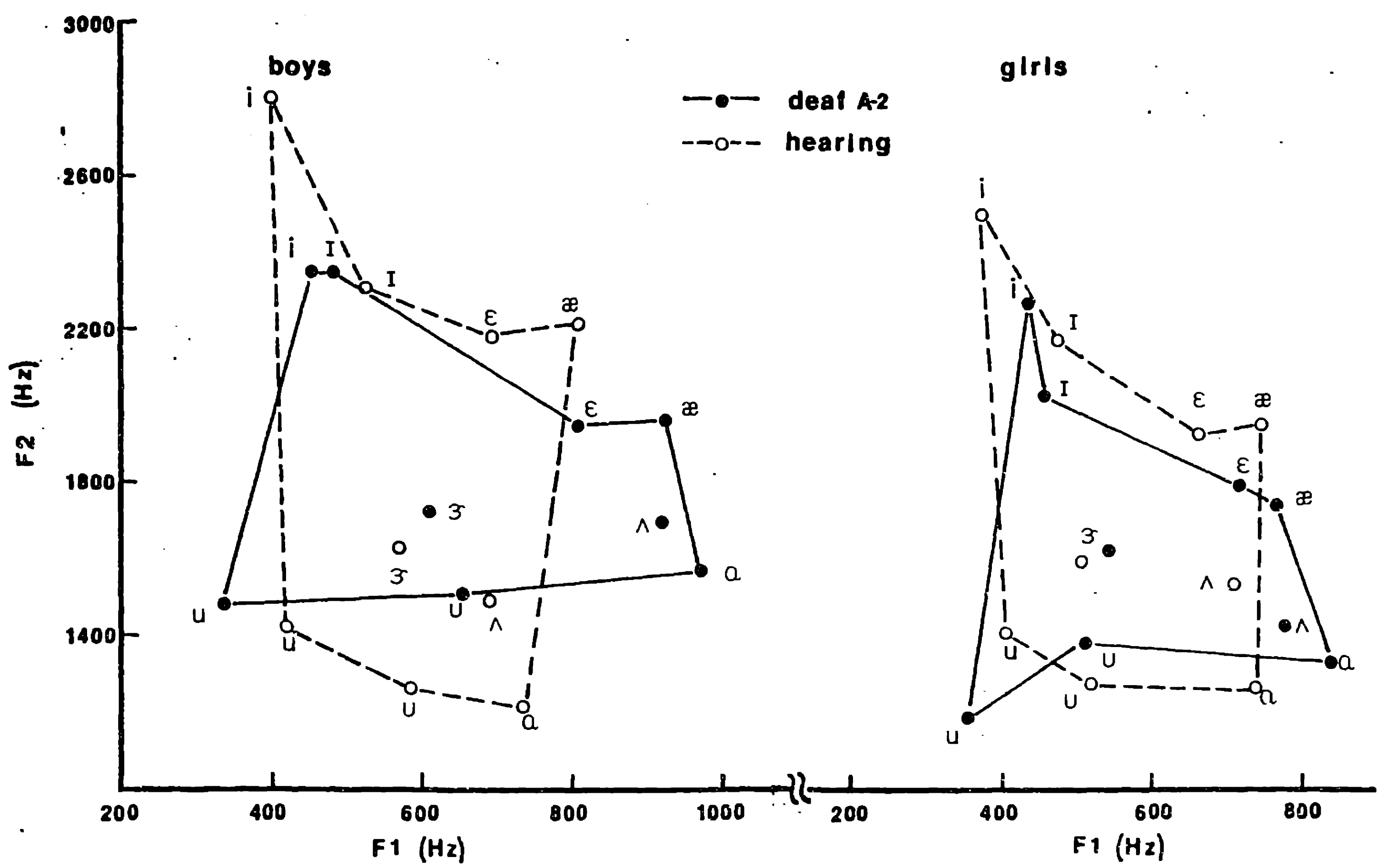


Figure 3.6 - Vowel diagrams (F2 versus F1) for the deaf speakers of Group A-2 and for the hearing controls. Formant data are plotted separately for boys (left) and for girls (right).

To some extent, the wider-than-normal range of F1 values observed for the speakers of Group A-2 may reflect measurement error, particularly in trying to estimate F1 for vowels with low F1 and high F0 (i.e., /i,u/). However, this exaggerated amount of F1 variation--in particular, the relatively high F1 for the vowels /ɑ/ and /æ/--may also be indicative of an exaggerated amount of jaw (and perhaps tongue) movement on the part of some of the deaf speakers. This latter possibility will be discussed more fully in Chapter 4.

The data presented in Figures 3.3 through 3.6 show that, for the majority of deaf speakers in this study, better articulatory skills (in terms of the range of F1 and F2 values used in producing the vowel targets) tend to be associated with greater vowel-to-vowel variability in F0. The scatter plot of Figure 3.7 suggests that greater F0 variability also tends to be associated with better speech intelligibility. For the fourteen speakers in Group A, a high positive correlation (r=.756) exists between percent intelligibility (as defined in Section 2.4.3) and vowel-to-vowel variability in F0 ("△F0").

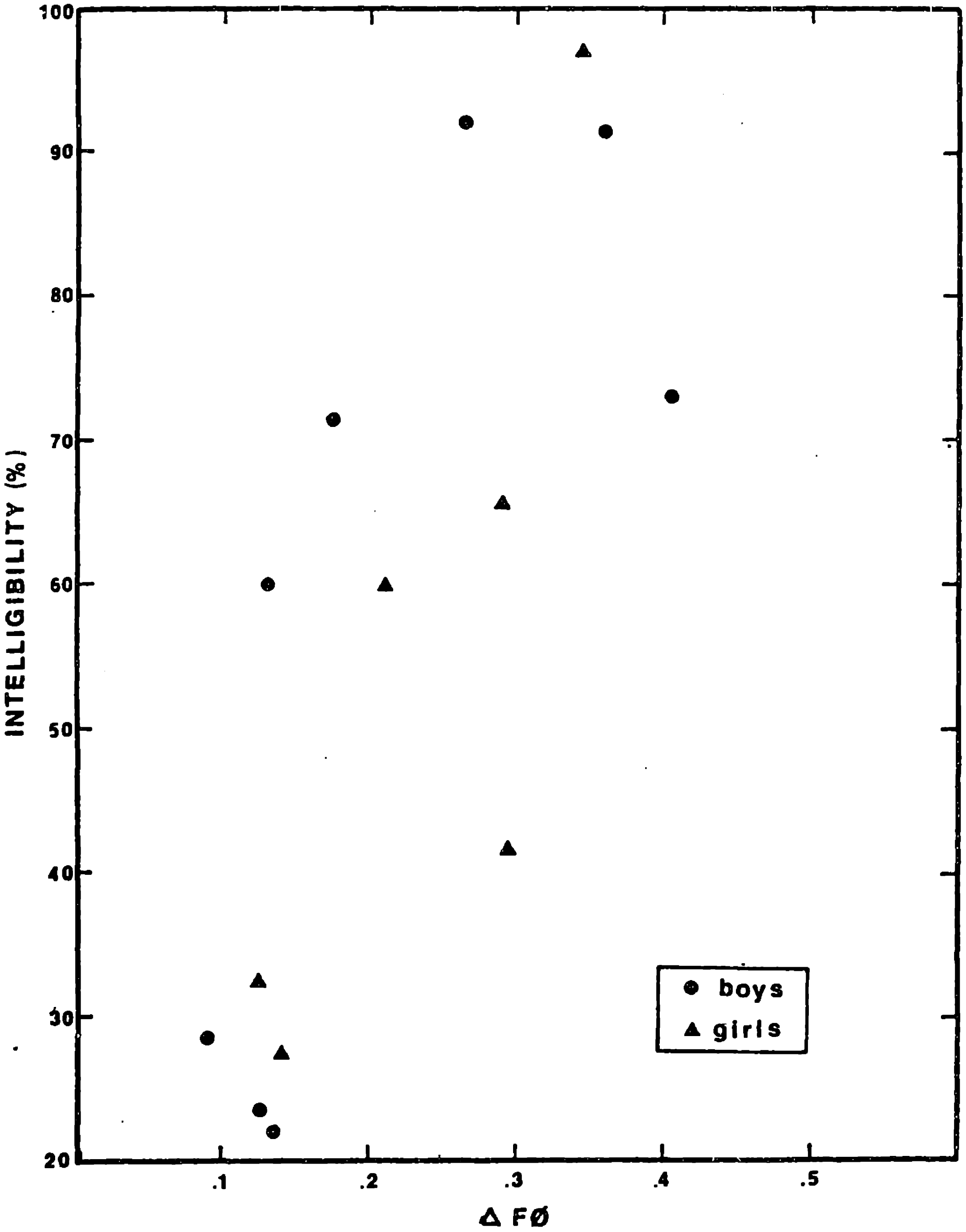


Figure 3.7 - Percent intelligibility versus vowel-to-vowel variability in F0 for the deaf speakers of Group A.

Group B:

(Speakers D14 and D16)

Returning to Figure 3.2, one can see that the two deaf girls in Group B approximate the hearing speakers with respect to both mean F0 and F0 variability. F0 (in Hz) is plotted as a function of vowel target for this group of girls and for the hearing controls in Figure 3.8. The two girls in Group B produce vowel-to-vowel variations in F0 that are slightly larger than those produced, on average, by the hearing speakers. However, comparison with Figure 3.4 shows that the F0 data for these two girls is much closer to normal than to the F0 values for the deaf speakers in Groups A-1 and A-2.

Figure 3.9 compares formant-frequency data for the girls in Group B with the corresponding data for the hearing girls. With the exception of the vowels /æ/ and /ʌ/, the mean F1 and F2 values produced by the two deaf girls are very close to those for the hearing controls, suggesting that their articulatory skills (at least with respect to vowel production) are also close to normal. Again, this observation is supported by the fact that intelligibility scores (Section 2.4.3) for these two girls are quite high (i.e., 95.0% and 86.3% respectively).

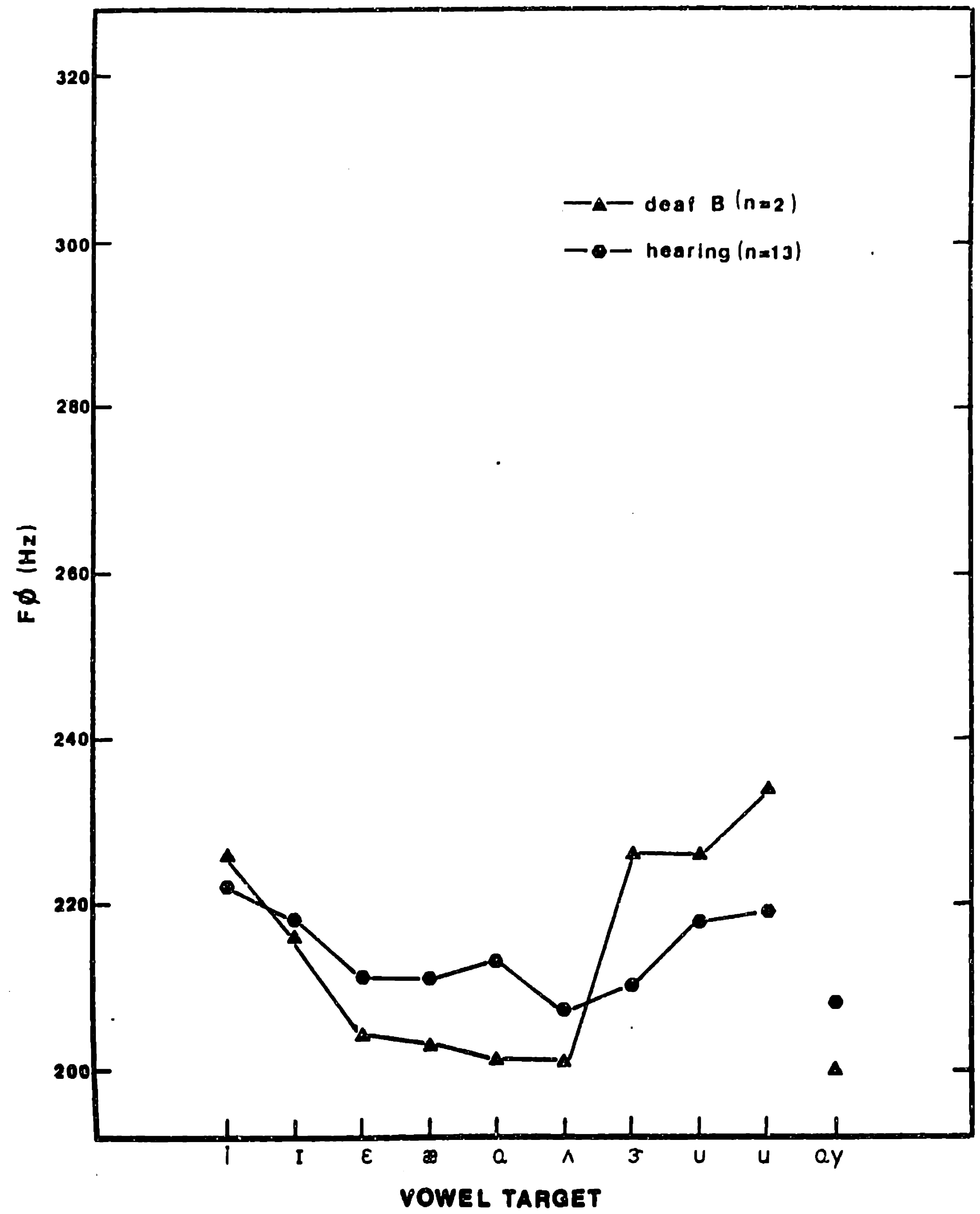


Figure 3.8 - Non-normalized FØ as a function of the vowel target for the deaf girls of Group B and for the hearing controls.

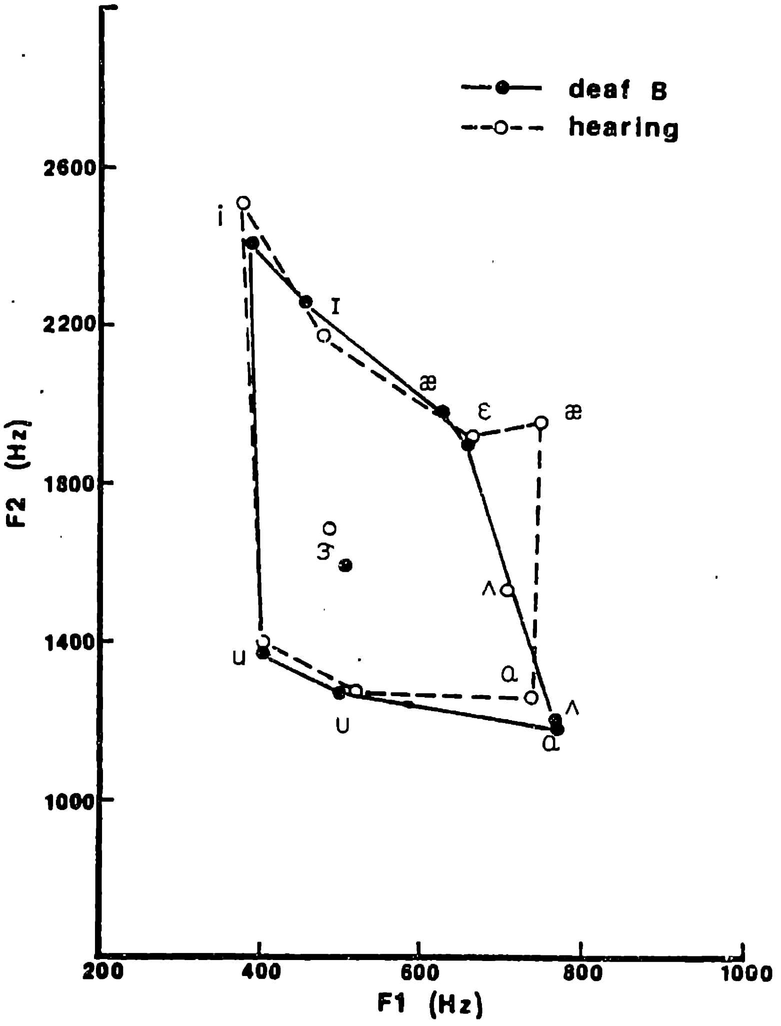


Figure 3.9 - Vowel diagrams (F2 versus F1) for the deaf girls of Group B and for the hearing girls.

Group C:

(Speakers D9-D12)

The third group of deaf speakers in Figure 3.2 (Group C) consists of four girls who produce an extreme amount of vowel-related variability in F0, particularly with respect to the hearing boys and girls but also with respect to the other deaf speakers. F0 (in Hz) is plotted as a function of vowel target for this group of girls in Figure 3.10. (For the sake of comparison, the F0 data in Figures 3.4, 3.8 and 3.10 are all plotted on the same scale.) The deaf girls in Group C tend to produce the vowel /ɑ/ with an F0 comparable to normal and the vowels /i/ and /u/ (especially /i/) with very high F0, resulting in a much greater range of cross-vowel fundamental frequencies than for any other group of speakers.

Formant-frequency data for the girls in Group C and for the hearing girls are compared in Figure 3.11. On average, the four deaf girls tend to produce all vowels with higher-than-normal first formants and to produce a substantially reduced range of F2 values. Examination of the individual data shown in Table 3.3, however, indicates that considerable differences exist among the deaf girls of Group C with respect to the range of F1 and F2 values which they produce, as well as with respect to mean F0 and speech intelligibility scores. Thus, statements about a relationship between F0 variability and articulatory skill for these four

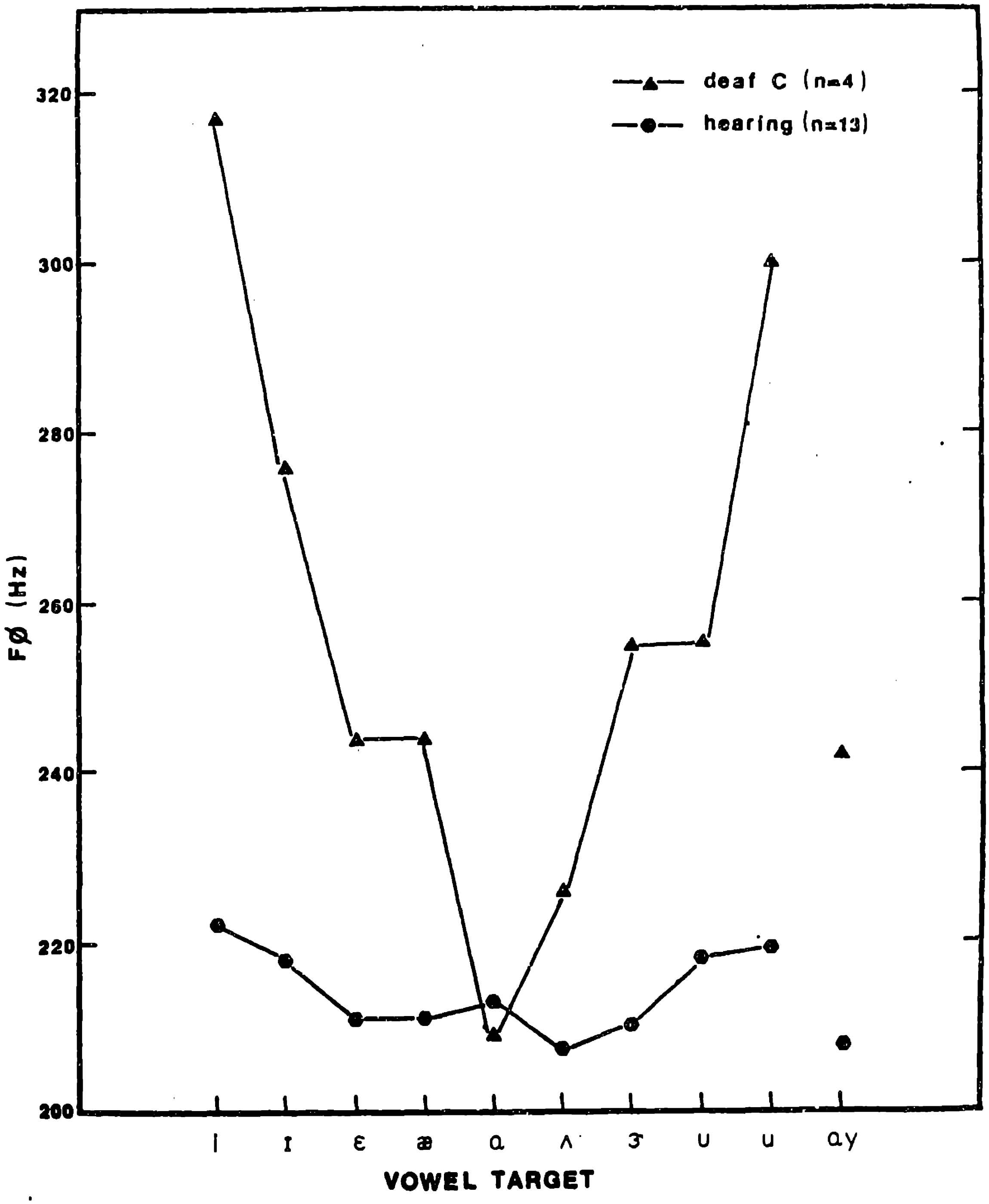


Figure 3.10 - Non-normalized F0 as a function of the vowel target for the deaf girls of Group C and for the hearing controls.

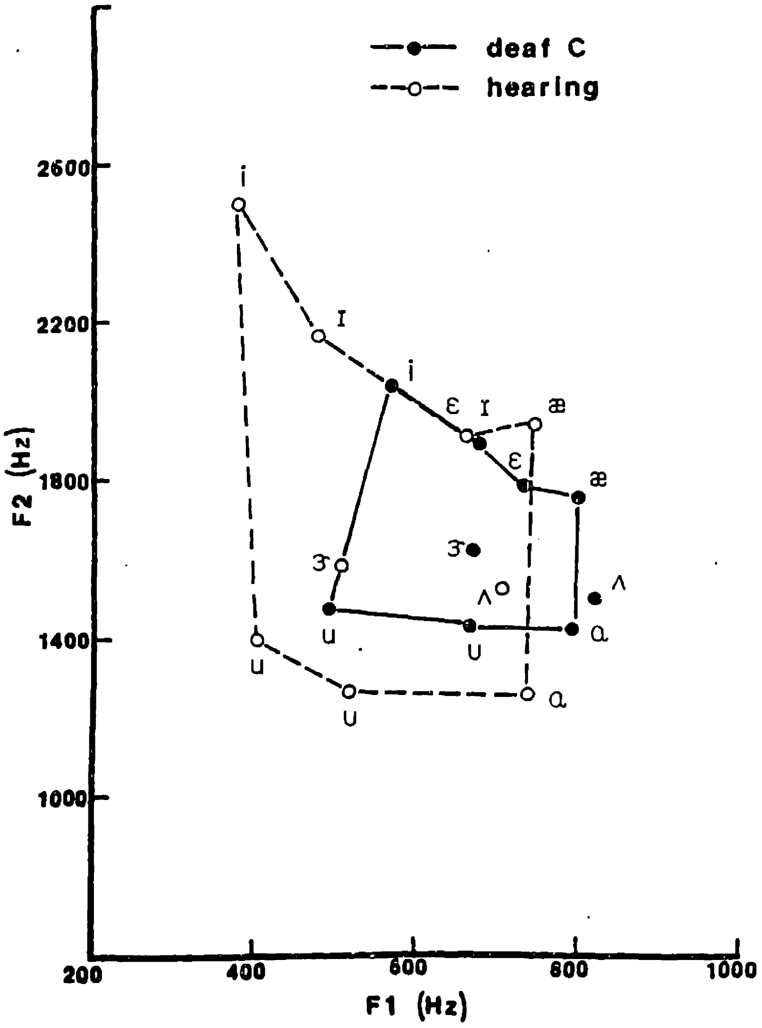


Figure 3.11 - Vowel diagrams (F2 versus F1) for the deaf girls of Group C and for the hearing girls.

Table 3.3 - Data for the Deaf Girls of Group C

speaker	D9	D10	D11	D12
△F0	.562	.520	.500	.490
mean F0 (Hz)	282	270	221	250
range F1 (Hz)	398	434	246	324
range F2 (Hz)	1180	667	656	518
intelligibility (%)	59.7	76.3	44.0	29.0
age (yr:mo)	13:10	14:2	13:11	13:11
hearing loss (dB)	83	107	107	112

girls, as a group, are impractical.

One statistic which is comparable for the four girls of Group C is age. As indicated in Table 3.3, all are between 13 years, 10 months and 14 years, 2 months old. Whether this association between chronological age and the use of excessive vowel-to-vowel variability in F0 is significant remains to be determined (e.g., by studying longitudinal F0 data); however, at least two considerations suggest that the association may not be coincidental. First, the four girls of Group C may have received comparable speech training at school, and it is conceivable that peculiarities of this training might be responsible for their difficulties with vowel-related F0 control.

Second, and perhaps more interesting, is the observation that the age in question, (i.e., approximately 14 years) falls within the adolescent growth spurt for females [Gol80]. While laryngeal growth during adolescence is much less pronounced for girls than for boys, anatomical studies show that significant increases in the size of the larynx occur in the circumpubertal years for both sexes [Kah75]. Furthermore, F0 data for hearing speakers incidate that a "voice change" (i.e., a lowering of F0) for females is typically noted between the ages of thirteen and fifteen [HolPau69].

Whether increased interactions between articulation and F0 are characteristic of the adolescent growth period (for male or for female speakers) is still open to question. However, if this were the case, it seems likely that compensation for such interactions would be more difficult for speakers lacking auditory feedback and, thus, that the resulting segmental variations in F0 would be relatively larger for deaf than for hearing boys and girls.

Another potential explanation for the excessive amount of F0 variability observed for the deaf girls of Group C is that they have learned to produce the vowel /ɑ/ in a normal manner (or at least with normal F0) but that their production of other vowels is more like that of the speakers in Group A-2. (Note that the F0 variability measure (i.e., "$\triangle$F0") for each speaker is based on measurements normalized relative to that speaker's mean F0 for the vowel /ɑ/. Since the speakers of Group C produce the vowel /ɑ/ with lower F0 than do the speakers of Group A-2, the normalized F0 measures for the other vowels and, thus, "$\triangle$F0" (which represents the range of these measures), will be larger in magnitude for the speakers of Group C.) The relatively poor articulatory capabilities of at least three of the girls in Group C (see Table 3.3), however, would seem to argue against this explanation.

3.2 Spectral Data

The relative amount of low-frequency energy in vowel spectra is dependent, in part, on the laryngeal posture--more specifically, the degree and abruptness of laryngeal closure--maintained during voicing. When the vocal folds are held in an abducted position, for example, complete closure of the glottis will not occur during voicing, and the glottal volume velocity waveform will be lacking in discontinuities. As a result, there will be relatively less high-frequency energy (or, alternatively, relatively more energy in the vicinity of the first and second harmonics) in the glottal source spectrum than in situations in which the vocal folds are approximated and come together along their entire length [Ste77]. (There may also be increased noise at high frequencies, due to turbulence at the glottis, and greater formant damping, due to increased losses in the glottis and trachea, when the vocal folds are in an abducted positon [Rot74];[Fis67].) Such incomplete closure of the glottis is thought to be associated with the production of "breathy" voice quality [Lad73];[Fis67], which, as noted in Chapter 1, is characteristic of the speech of some deaf individuals.

In order to obtain some information about the laryngeal postures maintained by the speakers in the present study (and of the relationship of these postures to F0 control), a measure was made of the relative amplitudes of the first formant and the first harmonic in the spectra of vowels with (relatively) high F1. The measure used can be defined as:

$$\frac{\Delta A}{\Delta F} = \frac{(AH - A0)}{(FH - F0)} \times 1000$$

where AH = amplitude of the highest harmonic in F1 (in dB)
A0 = amplitude of the first harmonic (in dB)
FH = frequency of the highest harmonic in F1 (in Hz)
F0 = frequency of the first harmonic (in Hz)

AH, A0, FH and F0 were measured from hardcopies of the DFT and LPC spectra (Section 2.4.3) for the vowels /ɛ,æ,ɑ,ʌ,ɝ,ɑy/, and the resulting $\Delta A/\Delta F$ values were averaged across vowels for each speaker. This averaged statistic will be referred to as "A/F".

As shown in Figure 3.12, each of the $\Delta A/\Delta F$ measures actually represents the <u>slope</u> of a line drawn from the peak of the first harmonic in the DFT spectrum for a given vowel target to the peak of the highest harmonic in F1. This frequency-normalized measure was used in order to account for cross-vowel and cross-speaker variations in the frequencies of F0 and F1, both of which can influence A0 and AH [Hou59]; [Fan56]. Measurments were made only for vowels in non-nasal consonantal contexts (since vowel nasalization can also influence the overall shape of the low-frequency energy

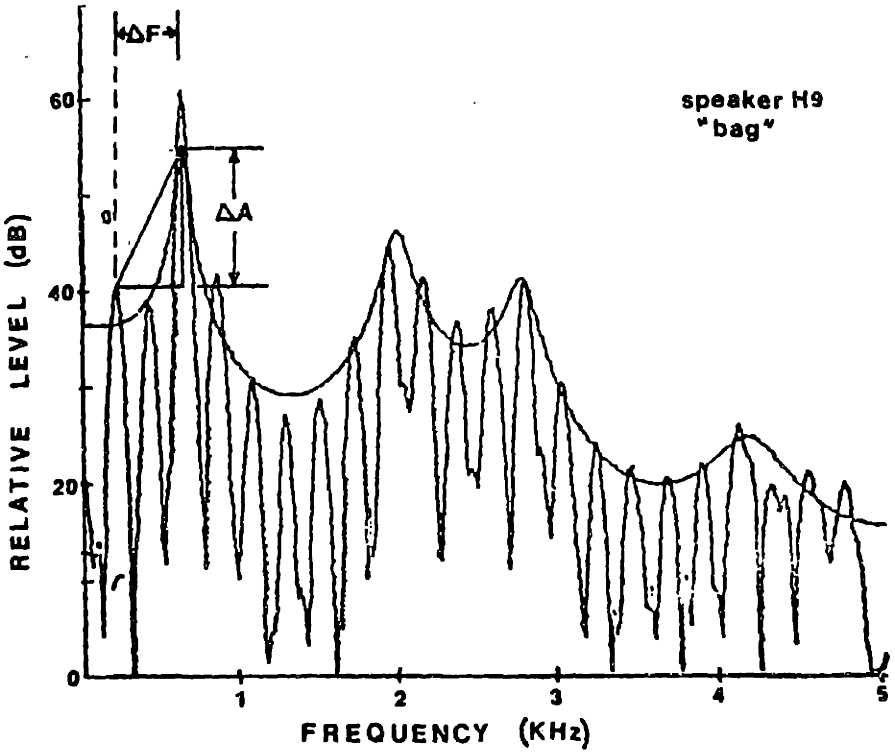


Figure 3.12 - Calculation of the ΔA/ΔF measure. (See text for explanation.)

Table 3.4 - Spectral Data

speaker	A/F (dB/KHz)	speaker	A/F (dB/KHz)
boys			
D1	10.6	H1	24.2
D2	37.4	H2	30.2
D3	41.6	H3	18.0
D4	31.0	H4	30.3
D5	27.4	H5	18.0
D6	7.3		
D7	18.7		
D8	18.3		
girls			
D9	34.5	H6	12.8
D10	40.6	H7	30.4
D11	22.7	H8	27.3
D12	27.8	H9	23.0
D13	30.4	H10	8.7
D14	13.6	H11	24.3
D15	39.6	H12	-0.8
D16	42.6	H13	19.3
D17	23.6	range	-0.8 - 30.4
D18	34.1		
D19	4.2		
D20	4.7		
range	4.2 - 42.6		

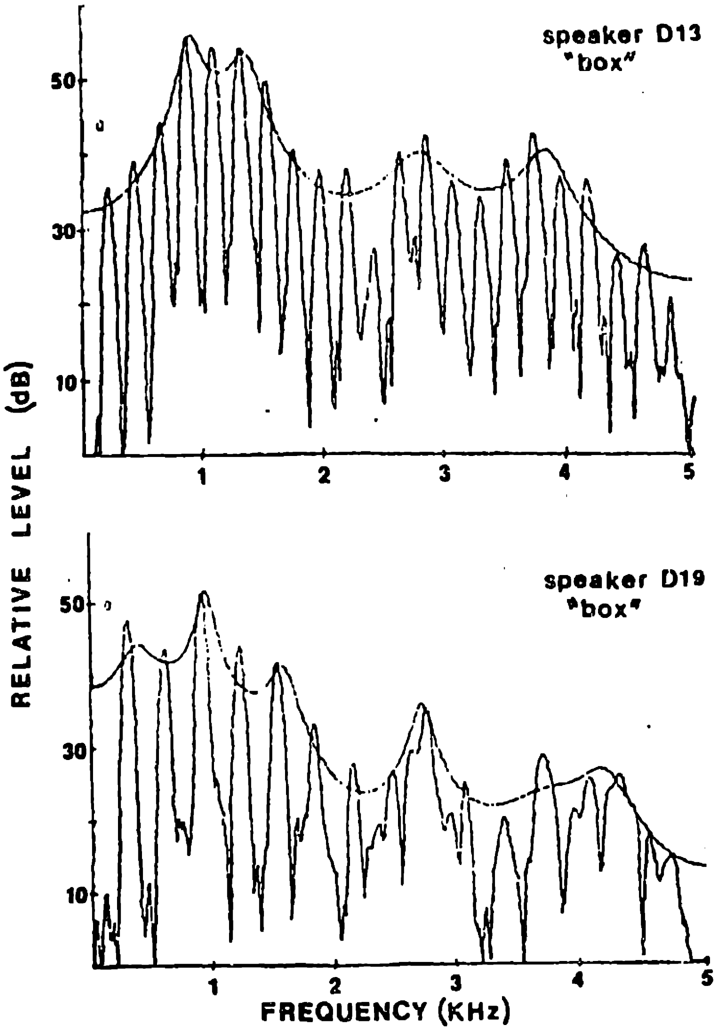


Figure 3.13 - Spectra characteristic of those produced by two deaf girls with high (top) and low (bottom) "A/F" scores. The spectra were computed near the center of the vowel target /ɑ/ in "box".

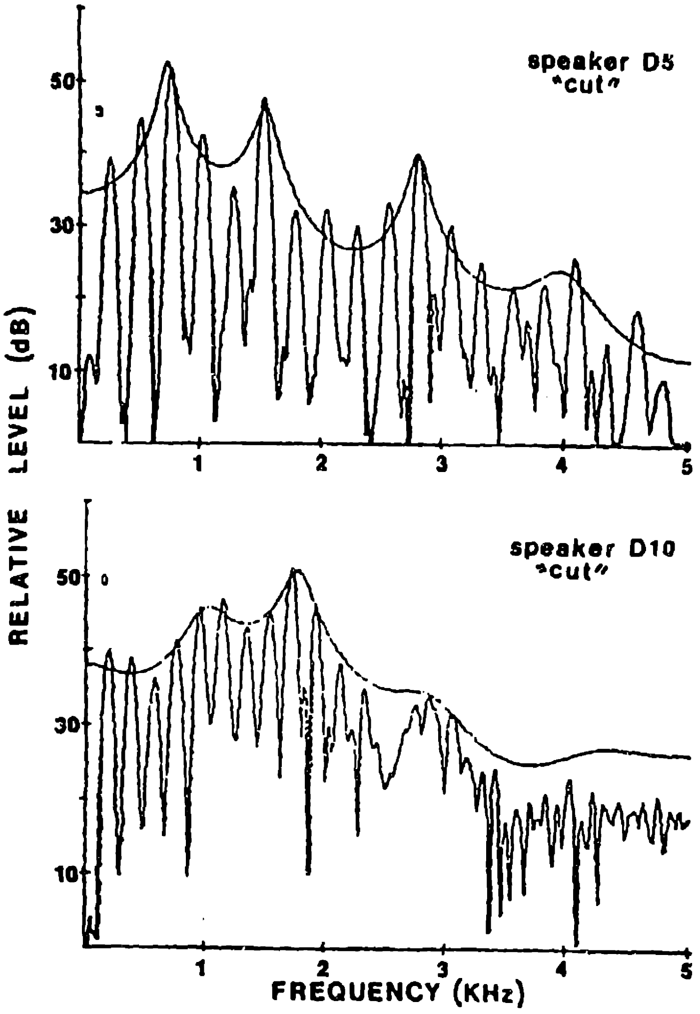


Figure 3.14 - Spectra characteristic of those produced by two deaf boys with high (top) and low (bottom) "A/F" scores. The spectra were computed near the center of the vowel target /ʌ/ in "cut".

spectrum [Fan60];[HouSte56]), and tokens for which the first harmonic was also the highest harmonic in F1 were eliminated.

Values of the averaged statistic "A/F" are listed for each of the deaf and hearing speakers in Table 3.4. These values vary across a relatively wide range for both groups of speakers: from a low of -0.8 dB/KHz to a high of 30.4 dB/KHz for the hearing boys and girls, and from 4.2 dB/KHz to 42.6 dB/KHz for the deaf speakers.

Spectra representative of those produced by deaf speakers with high and low "A/F" scores are shown in the top and bottom sections respectively of Figures 3.13 and 3.14. Examination of these figures shows that the spectra associated with low "A/F" are characterized not only by a fundamental that is higher in amplitude (relative to the first formant) but also by a considerable amount of noise in high-frequency harmonics. As noted earlier, the presence of such noise would be consistent with the generation of turbulent airflow at the (partially-opened) glottis [Rot74];[Fis67]. (Since the spectra were computed over a number of pitch periods, a rapidly changing or unsteady F0 might also contribute to the appearance of noise at high frequencies.)

The voices of the two deaf girls with the lowest "A/F" scores (speakers D19 and D20) were described as "breathy" by the linguists performing the phonetic transcriptions, while speaker D11 was described as "breathy" by one transcriber and "strained" or "harsh" by another. Judging by the "A/F" scores (which were often higher than those for any of the hearing controls) and by the transcribers' comments, however, breathiness did not appear to be a problem for the majority of deaf speakers in the study.

As can be seen in the scatter plot of Figure 3.15, there is a tendency among both the deaf and hearing speakers for "A/F" to increase with the amount of vowel-related F0 variability used (i.e., "ΔF0"). Hence, it is possible that the relatively high "A/F" scores noted for some of the deaf speakers, to some extent, represent a bias in this particular spectral measure (since "ΔF0" is also high for many of the deaf speakers). It is also, possible, however, that these high values of "A/F" are indicative of a mode of vocal-fold vibration which is somewhat more "pressed" or "tense" than normal. Several investigators have noted, for example, that deaf speakers often maintain a laryngeal posture in which the vocal folds are too tightly approximated [Pet46];[Hud37], and inverse filtering experiments (with the speech of hearing individuals) have shown that the glottal source spectrum for this type of phonation is characterized by a relatively weak fundamental with strong higher-frequency harmonics [SunGau78].

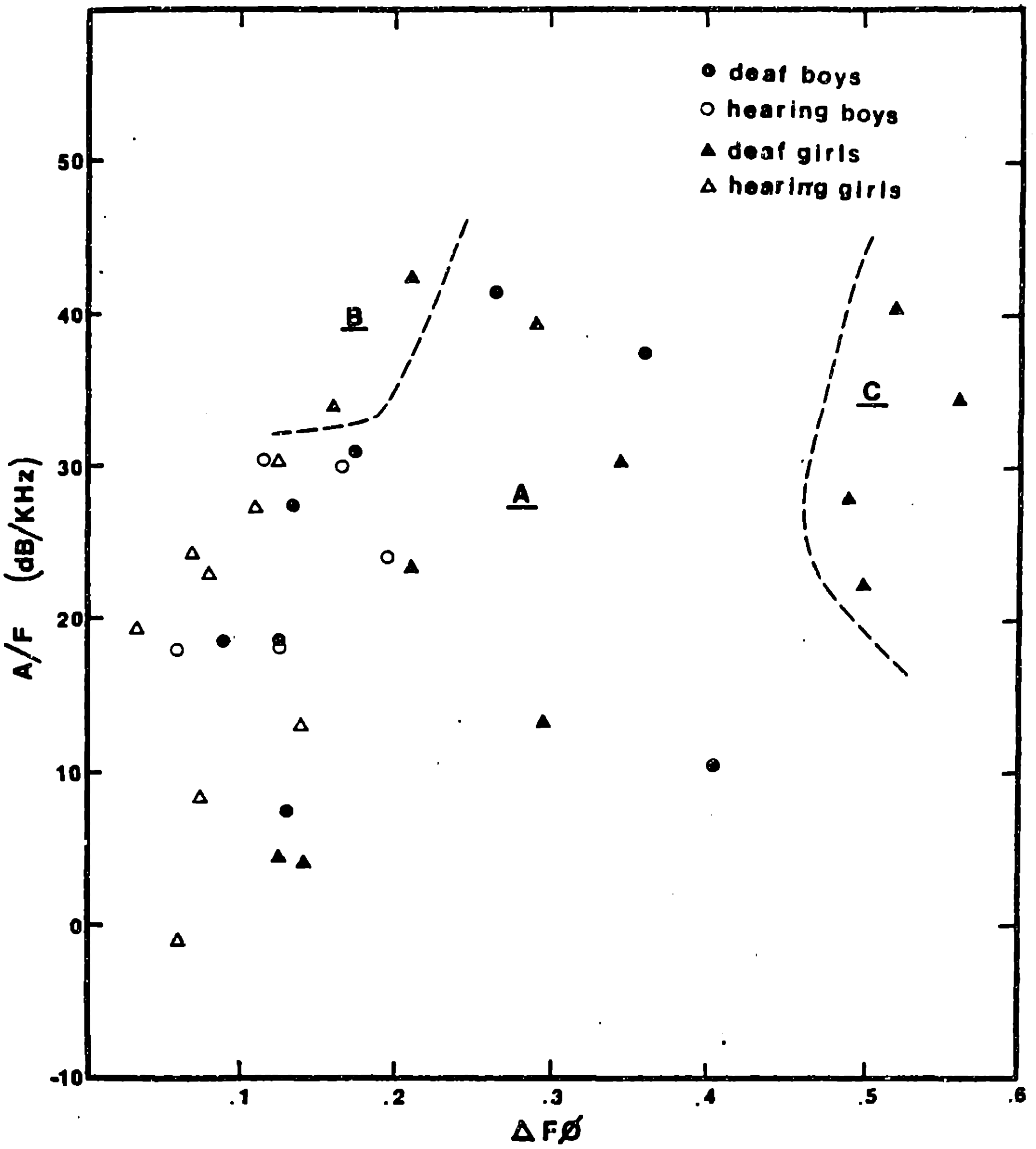


Figure 3.15 - Scatter plot of the spectral measure "A/F" versus vowel-to-vowel variability in F0 for the deaf and hearing speakers. (See text for an explanation of Groups A, B and C.)

Despite the overall tendency for "A/F" to increase with "$\Delta F0$", the relationship between these two variables is not particularly consistent. For example, of the five speakers in Group A with the lowest "A/F" scores (Figure 3.15), three use relatively little F0 variability while the remaining two speakers produce large vowel-to-vowel variations in F0. The correspondence between "A/F" and "mean F0" is also not very close (Figure 3.16) . Among the deaf speakers, "A/F" tends to decrease somewhat with increasing F0, but, again, their is considerable scatter in the data. Furthermore, there does not seem to be a consistent division among groups of speakers (i.e., A,B,C) with respect to the "A/F" score.

Table 3.5 shows partial correlation coefficients (for the deaf speakers) computed between "A/F" and the various other statistics considered in this chapter. As suggested by the scatter plots of Figures 3.15 and 3.16, correlations between "A/F" and "$\Delta F0$" and between "A/F" and "mean F0" are not particularly high; nor are those between "A/F" and measures indicative of the deaf speakers' articulatory skills (i.e., "range F1" and "range F2"). These observations suggest that, for the majority of the deaf boys and girls, the positioning, or spread, of the vocal folds (at least as evidenced by the spectral data) may be relatively independent of overall laryngeal tension and of vowel articulation.

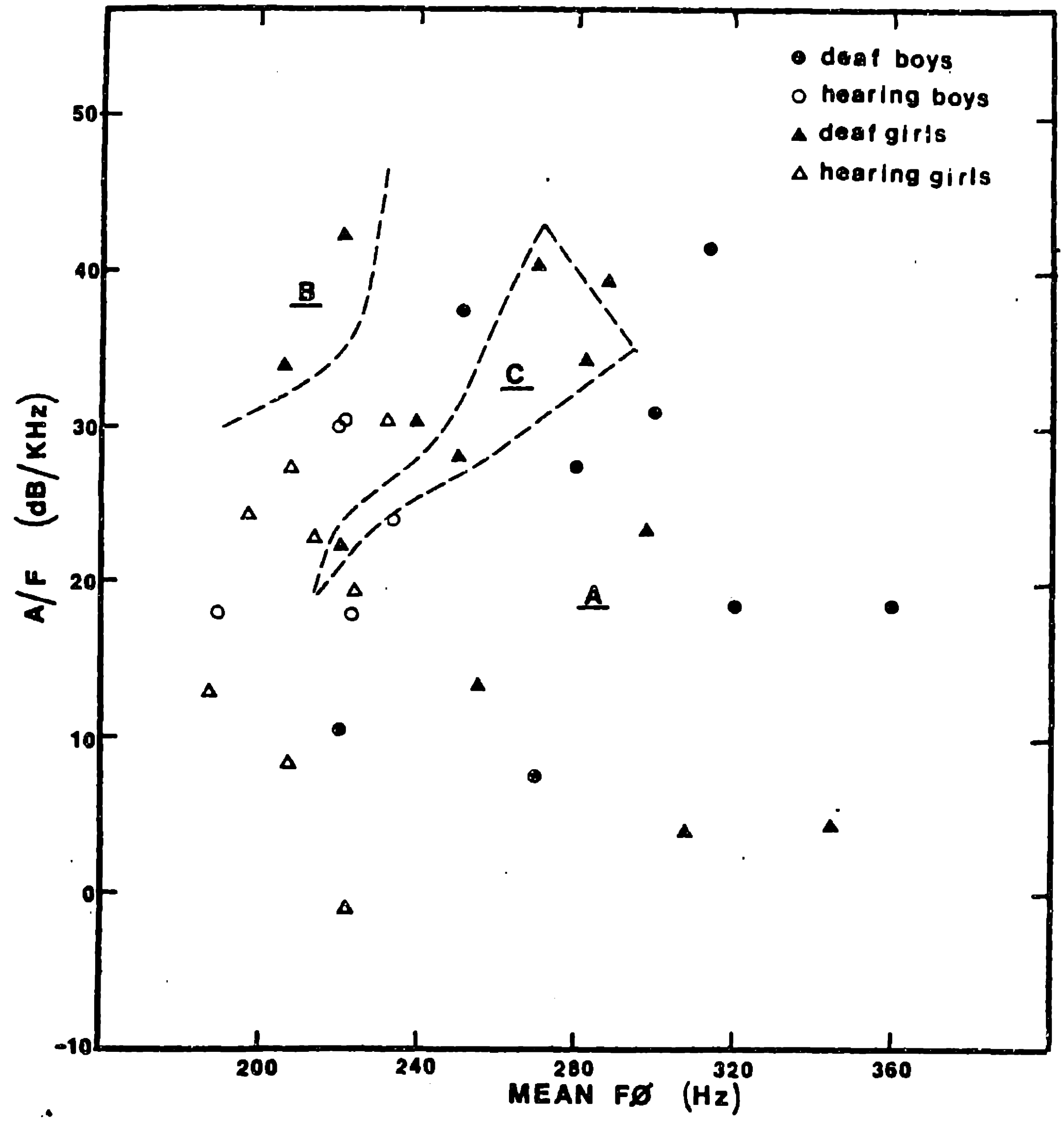


Figure 3.16 - Scatter plot of the spectral measure "A/F" versus mean F0 for the deaf and hearing speakers. (See text for an explanation of Groups A, B and C.)

Table 3.5 - Partial Correlations between "A/F" and Other Variables for the Twenty Deaf Speakers.

A/F vs.	r
mean F0	-.286
△F0	.376
range F1	.186
range F2	.472
intelligibility	.628

Somewhat surprisingly, given the relatively low correlations between "A/F" and the other variables, a fairly strong correlation (r=.628) exists between "A/F" and speech intelligibility. One possible explanation for this finding is that variations in the frequency of F1 (and perhaps F2) may be somewhat less salient perceptually when the low-frequency energy spectrum is "flattened" by increased losses in the glottis or subglottal system, by the introduction of extraneous formants (due to coupling to the subglottal or nasal systems) or by a high-amplitude first harmonic [Ste-pc]. Thus, a low "A/F" score might be associated with low speech intelligibility, even in cases in which a relatively wide range of vowel sounds (i.e., of first and second formant frequencies) is produced.

3.3 Hearing Loss

In Section 3.1.3, the suggestion was made that chronological age might contribute to observed differences in the amount of (vowel-related) F0 variability used by deaf speakers. Another variable which might influence F0 control (either directly or indirectly) is severity of hearing loss. Figure 3.17 shows a scatter plot of hearing loss versus vowel-to-vowel variability in F0 ("ΔF0") for the twenty deaf boys and girls. (The measure of hearing loss used is defined in Table 2.1.) Observations to be made from this plot include the following:

1) For the fourteen boys and girls in <u>Group A</u>, a negative, though not particulary high, correlation (r=-.344) exists between hearing loss and F0 variability. On average, the hearing losses of those speakers producing large vowel-to-vowel variations in F0 (Group A-2) are approximately 8 dB less than those of the speakers producing smaller variations in F0 (Group A-1) (i.e., 100 dB for Group A-2 as compared with 108 dB for Group A-1).

2) The two girls in <u>Group B</u> (i.e., those girls who are most like the hearing speakers with respect to mean F0, F0 variability and formant variability) have hearing losses that are less severe than those of most of the other deaf speakers--83dB and 90dB as compared with an average of 102 dB for all twenty speakers.

3) The hearing losses of three of the four girls in <u>Group C</u> (i.e., those speakers producing the greatest amounts of vowel-related F0 variability) are quite severe; however, that of the fourth girl is lower than those of all but one of the other nineteen speakers.

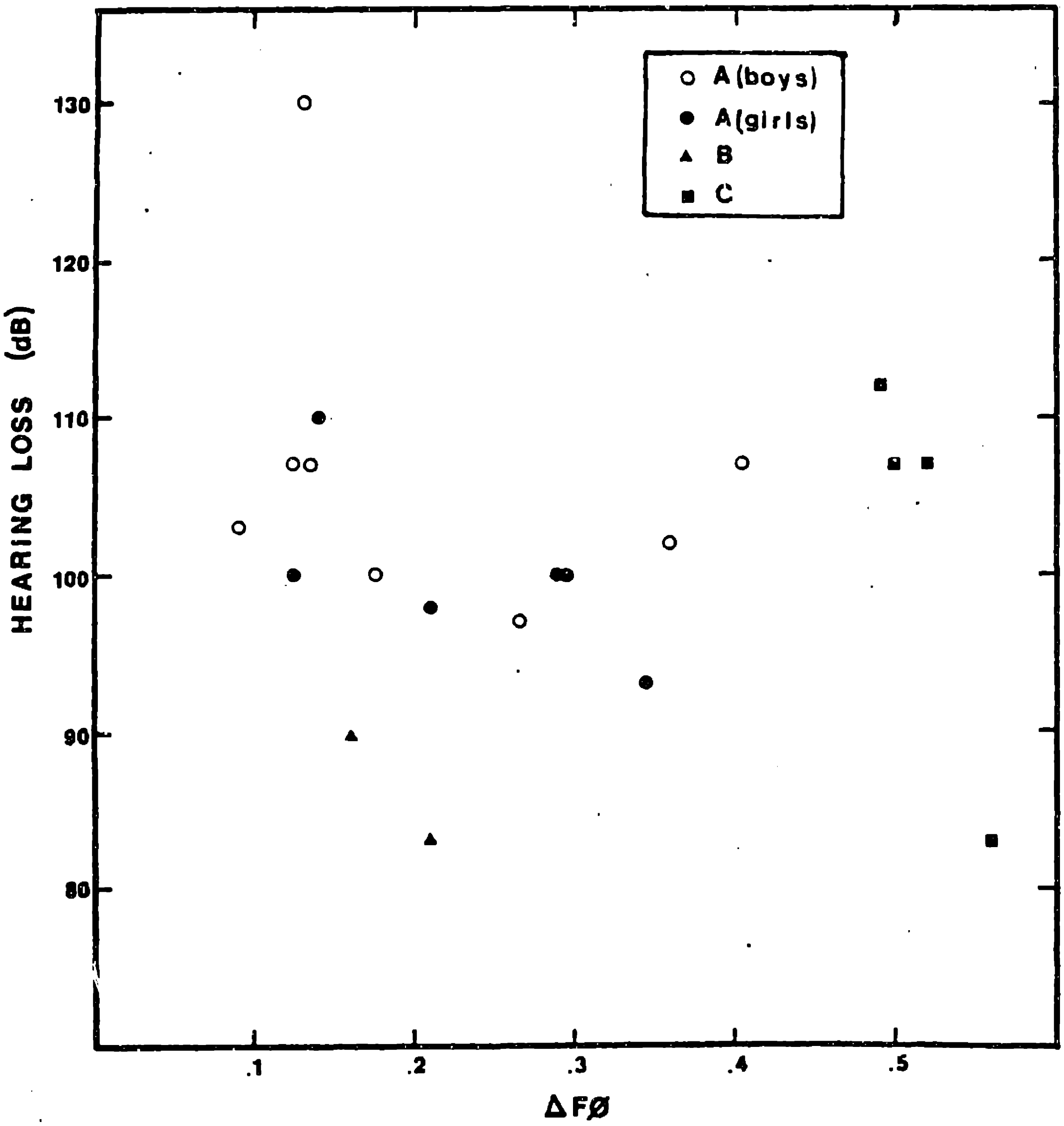


Figure 3.17 - Scatter plot of hearing loss versus vowel-to-vowel variability in F0 for the deaf boys and girls.

When the twenty deaf speakers are treated as a single group and partial correlation coefficients are computed between hearing loss and the various (measured) statistics described in this chapter, the highest correlation is between hearing loss and the range of F2 values produced (r=-.708) (see Table 3.6). This inverse relationship between severity of hearing loss and range of F2 is illustrated schematically in the scatter plot of Figure 3.18. Here it can be seen that, regardless of group affiliation (i.e., A,B,C), those deaf speakers whose hearing losses are more severe tend to produce a much narrower range of F2 values than do those speakers whose hearing losses are lower.

Examination of Table 3.6 shows that an inverse correlation also exists between severity of hearing loss and the range of F1 values produced, but the correlation coefficient in this case is considerably smaller in magnitude (r=-.288). Correlations between hearing loss and mean F0 and between hearing loss and F0 variability are relatively low, approximately equal in magnitude and opposite in sign (r=.145 and r=-.144) respectively)--observations which are not surprising given the differences in the F0 data for the three groups of speakers and the relationships between mean F0 and F0 variability described in Section 3.1.3.

Table 3.6 - Partial Correlations between Hearing Loss and Other Variables for the Twenty Deaf Speakers.

Hearing Loss vs.	r
mean F0	.145
F0	-.144
range F1	-.288
range F2	-.708
A/F	-.555

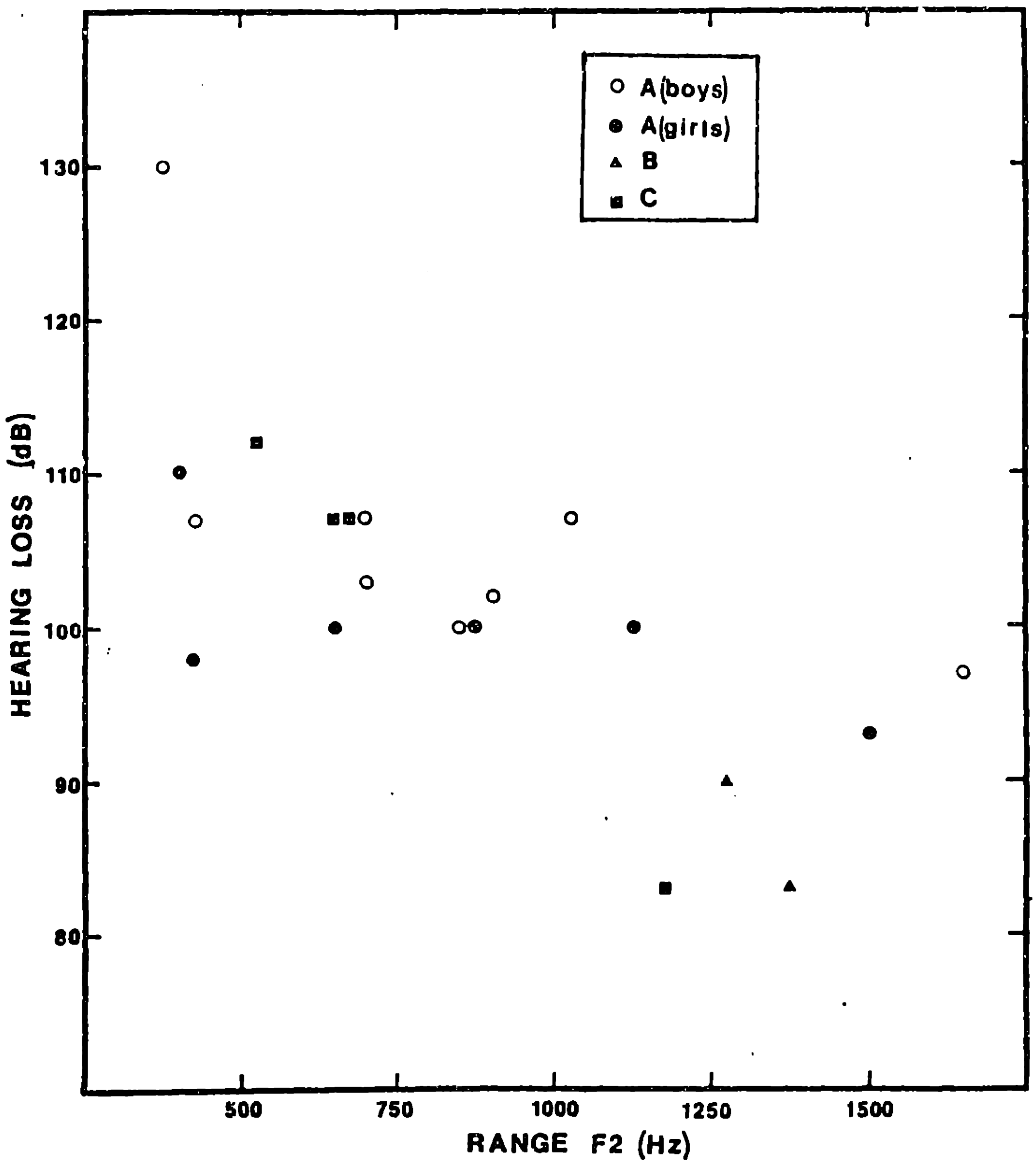


Figure 3.18 - Scatter plot of hearing loss versus range F2 for the deaf boys and girls.

Finally, a negative correlation exists between hearing loss and "A/F", suggesting that "breathy" voice quality (at least to the extent to which it is evidenced by enhanced low-frequency spectral energy) may be more common among deaf speakers whose hearing losses are relatively more severe. While this possibility needs to be examined more carefully, it would seem reasonable, in that any residual hearing or tactile response such speakers might have is likely be at very low frequencies [BBN75]. Hence, the production of vowels with an overly high-amplitude first harmonic (relative to the amplitudes of higher frequency harmonics) might represent an attempt by a deaf speaker to monitor (i.e., to hear or to feel) his or her own speech.

3.4 Summary and Discussion

The major observations to be made from the data presented in this chapter are the following:

1) Large individual differences exist among the deaf boys and girls with respect to mean F0, vowel-related variability in F0 and the range of F1 and F2 values produced. A very small number of the deaf speakers in the study (2 of 20) approximate the behavior of the hearing speakers with respect to all variables measured;

intelligibility scores for these speakers are high and their hearing losses are less severe than those of most of the other deaf speakers.

For the majority of deaf speakers (14 of 20), an inverse relationship exists between mean F0 and F0 variability. Those speakers who use greater F0 variability (and a lower mean F0) tend to produce a wider range of F1 and F2 values, to be more intelligible and to have hearing losses that are relatively less severe than do those speakers who use less F0 variability (and a higher mean F0).

The remaining deaf speakers (4 of 20) use an excessive amount of vowel-related variability in F0 relative to both the hearing speakers and to the other deaf boys and girls. While large individual differences exist within this group with respect to mean F0, formant variability, intelligibility and severity of hearing loss, the four speakers are approximately the same age.

2) Although the amount of F0 variability used by the deaf speakers varies across a wide range, the direction in which F0 varies as a function of vowel height is the same for the deaf and hearing speakers. This observation holds both on average

and when the deaf speakers are divided into groups. In all cases, F0 tends to be higher for high vowels, such as /i/ and /u/, than for low vowels, such as /ɑ/.

3) While the relative amount of low-frequency energy in the vowel spectra varies considerably among the deaf speakers, this variability is not consistently related to mean F0, F0 variability or to the range of F1 and F2 values produced. Nonetheless, better speech intelligibility appears to be associated with the production of vowels in which the amplitude of the first formant is high relative to that of the fundamental.

4) Correlations between severity of hearing loss and measures indicative of articulatory skill (i.e., range F1 and range F2) are higher than between hearing loss and measures indicative of F0 control (i.e., mean F0 and F0 variability).

On the basis of these observations, a number of statements can be made about the relationship between laryngeal control and proficiency of vowel articulation in the speech of deaf individuals (or at least for the deaf boys and girls participating in this study). First, given the amount of variability across deaf speakers with respect to all parameters measured, it would seem unwise to make generalizations about "the speech of the deaf" solely on the basis of averaged data. This caution would presumably also apply to relationships among variables other than those investigated in the present study.

Second, the observation that greater vowel-to-vowel variability in F0 tends to be associated with greater formant variability and with better speech intelligibility for the majority of deaf boys and girls would seem to refute the hypothesis that deaf speakers use exaggerated F0 variability as a substitute for articulatory variability in order to distinguish one vowel sound from another [Hor77];[AngKop64]. Instead, the similarities in the shapes of the F0 vs. vowel target curves for the deaf and hearing speakers suggest that the vowel-related variations in F0 produced by these two groups are *qualitatively* similar and, in particular, are a *consequence* of articulatory maneuvers used in vowel production. Mechanisms which might be responsible for these interactions between vowel articulation and F0 will be considered in the following chapter.

Third, the lack of a consistent relationship between the spectral measure "A/F" and the measure of F0 variability " F0" suggests that vocal-fold posture (i.e., vocal-fold spread) is not a major factor in determining the exaggerated vowel-related variations in F0 produced by many of the deaf speakers. The implications of this finding for the (hypothetical) mechanisms for F0 change discussed in Section 1.3 will also be considered in the following chapter.

Finally, the observation that severity of hearing loss correlates more highly with measures of formant variability than with F0 variability and mean F0 suggests the possibility that hearing loss may have only a secondary influence on fundamental-frequency control, whereas the primary influence is on a deaf speaker's articulatory skills. This observation, plus the apparent interdependence of most of the variables examined in this chapter, suggest the following interpretation of the data. (Note that this interpretation may be most applicable to deaf speakers attending a school, like the Clarke School, in which a significant amount of attention is devoted to speech training.)

When hearing loss is less severe (than some criterion yet to be determined), vowel articulation and mean F0 will tend to be close to normal. Intelligibility will be high and vowel-to-vowel variations in F0 will tend to be comparable to or slightly greater than normal. As hearing loss becomes more severe, articulatory proficiency will tend to deteriorate and mean F0 will tend to increase relative to normal. This increase in mean F0 may be a consequence of changes in articulation (e.g., a more rigid or tense overall posture of the speech-generating structures may influence both mean F0 and formant variability) or of a relatively independent change in laryngeal tension.

In any event, the amount of vowel-to-vowel variability in F0 produced by a deaf speaker will be influenced by the way in which his/her articulatory strategies, and perhaps articulatory posture or laryngeal tension, change relative to normal. Those speakers who use relatively little articulatory variability (and who maintain an articulatory posture or laryngeal tension which leads to high mean F0) will produce relatively small vowel-related variations in F0. On the other hand, those speakers who tend to exaggerate their articulation (e.g., by using an exaggerated amount of jaw movement or by associating an inappropriately high overall tension with certain vowel sounds) will tend to produce larger-than-normal vowel-related variations in F0. Finally, these various trends can be overidden by other factors, such as age (i.e., puberty) and peculiarities of speech training.

CHAPTER 4

MECHANISMS FOR F0 CHANGE

The objective of this chapter is to identify mechanisms which might account for the (exaggerated) vowel-to-vowel variations in F0 produced by the deaf boys and girls. This objective is pursued through: 1) an examination of the influence of consonantal context on the F0 contours produced by the deaf and hearing speakers; and 2) an interpretation of these F0 data, together with the data from the preceding chapter, in terms of the hypothetical mechanisms for vowel-related F0 change discussed in Section 1.3.

4.1 The Influence of Consonantal Context on F0

4.1.1 Introduction

The analyses described in the following sections were guided by three assumptions:

1) If acoustic coupling between F0 and F1 plays a major role

in determining the segmental variations in F0 produced by the deaf speakers, an inverse relationship should exist between the frequency of the first formant and that of the fundamental for consonants as well as for vowels. In particular, F0 should be relatively high on <u>nasal</u> consonants (and perhaps on the sonorants /r,l,w/), independent of their place of articulation or of the height of adjacent vowel sounds. One might also expect changes in the F0 contour associated with the introduction of a low-frequency first formant (e.g., at the boundary between a low vowel and a nasal consonant) to be relatively abrupt, if F0 and F1 were sufficiently close in frequency and vocal-tract damping were sufficiently low. (See Section 1.3 for a discussion of factors motivating this latter expectation.)

2) Information about the relative influence of jaw position and tongue height on the deaf speakers' control of F0 might be obtained by examining the way in which F0 varies over the course of syllables beginning with <u>labial</u> (stop) consonants. During the production of such syllables, the position of the jaw is constrained at syllable onset, while the tongue body is (relatively) free to anticipate a configuration appropriate to that of the following vowel [Per69]. If jaw position is the major factor controlling F0, F0 should be relatively constant near syllable onset

for syllables with initial labials, independent of vowel context, but should change significantly over the course of syllables in which the jaw is lowered for the production of a low vowel. On the other hand, if tongue height is crucial in determining F0, coarticulatory changes in the height of the tongue body should be reflected in the F0 data (i.e., F0 near syllable onset for labials should be vowel dependent).

3) Additional information about the influence of tongue position on F0 might be obtained by examining F0 as a function of <u>place of consonant articulation</u>. If raising the tongue body plays a major role in increasing F0, then F0 should be relatively higher near the onset of syllables beginning with velar (stop) consonants (for which the tongue body is the primary articulating structure [Per69]) than for syllables beginning with labials or alveolars, and F0 change should be greatest over the course of syllables in which velar consonants are followed by low vowels.

As noted earlier, two sets of (narrow) phonetic transcriptions were obtained for each of the deaf boys and girls. These transcriptions were used, together with wideband sound spectrograms, in deciding whether or not an "acceptable" version of a given consonantal phoneme had been produced (i.e., whether or not the segmental feature(s) of interest were present). The specific criteria used for each class of consonants and the number (and/or percentage) of target phonemes meeting these criteria are provided in the sections which follow.

Because the test phrases were repeated only once by each boy and girl, a maximum of twenty repetitions of a given target noun was possible for the deaf speakers. With such a limited amount of data, it was impractical to examine interspeaker differences for individual target utterances. Within each class of consonants, however, F0 data for sets of target nouns (e.g, nouns containing high versus low vowels) were compared for the deaf speakers in Groups A-1, A-2, B and C (i.e., the groups defined in the preceding chapter) and for the hearing controls. The F0 contours were also examined qualitatively in order to identify differences in the types of contours produced which might be important.

In the following four sections (Sections 4.1.2 through 4.1.5), F0 data are presented graphically, and typical (and atypical) F0 contours for each consonantal context are described. The major trends in these data are also interpreted relative to the three assumptions outlined above. (A more complete discussion of mechanisms for F0 change, however, is postponed to Sections 4.2.1 and 4.2.2.) Numerical F0 data, together with the number of tokens upon which each F0 measurement is based, are provided in Appendix A2.

4.1.2 Nasal Consonants

Three criteria were used in deciding whether an "acceptable" nasal consonant had been produced by a deaf speaker: 1) evidence of nasalization on a wideband sound spectrogram of the utterance; 2) agreement between the two transcribers on the intended place of consonant articulation; and 3) identification of the consonant as a nasal by at least one of the two transcribers.

In the case of the consonant /m/ in syllable-initial position, a total of eighty tokens was possible for the deaf speakers and, of these, sixty-nine (86.2%) met the above criteria. Five of the sixty-nine consonants which were accepted were "ambiguous" in that they were identified as nasals (or as nasals coarticulated with stops) by one of the two transcribers and as stops (i.e., /b/'s) by the other. For syllable-initial /n/'s, forty tokens were possible; of these, twenty-eight (70%) were accepted and one was ambiguous. (See Tables A2.1 and A2.2 in Appendix A2 for the number of tokens accepted for each target noun.)

Figures 4.1 and 4.2 compare F0 data for the deaf and hearing speakers for target nouns beginning with /m/ and /n/ respectively. (These F0 data are based on measurements made at the "centers" of the nasal consonants and at the "centers" of the vowel targets, as described in Section 2.4.1). For the deaf group of speakers, F0 on the labial consonant /m/ appears to be closely dependent on the F0 value at the center of the following vowel (Figure 4.1). F0 on the /m/ is comparable to (or slightly lower than) F0 on the vowel for all nouns except "mitt", for which F0 on the vowel is considerably higher.

For the hearing group of speakers, on the other hand, F0 remains relatively constant on the consonant /m/, independent of vowel context, while changing in the expected manner on the vowel targets (i.e., F0 is higher for high than for low vowels). F0 is higher on the /m/ than on the vowel in nouns containing low vowel targets (i.e., "man" and "mug"); for nouns containing high vowels, the difference in F0 between consonant and vowel is negligible.

In the case of nouns beginning with the alveolar consonant /n/, F0 on the consonant is higher than (or comparable to) F0 on the vowel for both groups of speakers (Figure 4.2), and cross-vowel changes in F0 (on both the consonant and the vowel) are slightly greater in magnitude for the hearing boys and girls.

For nasal consonants in syllable-final position, a total of one hundred and twenty tokens was possible, and ninety-four (78.3%) met the criteria for acceptance defined above. With

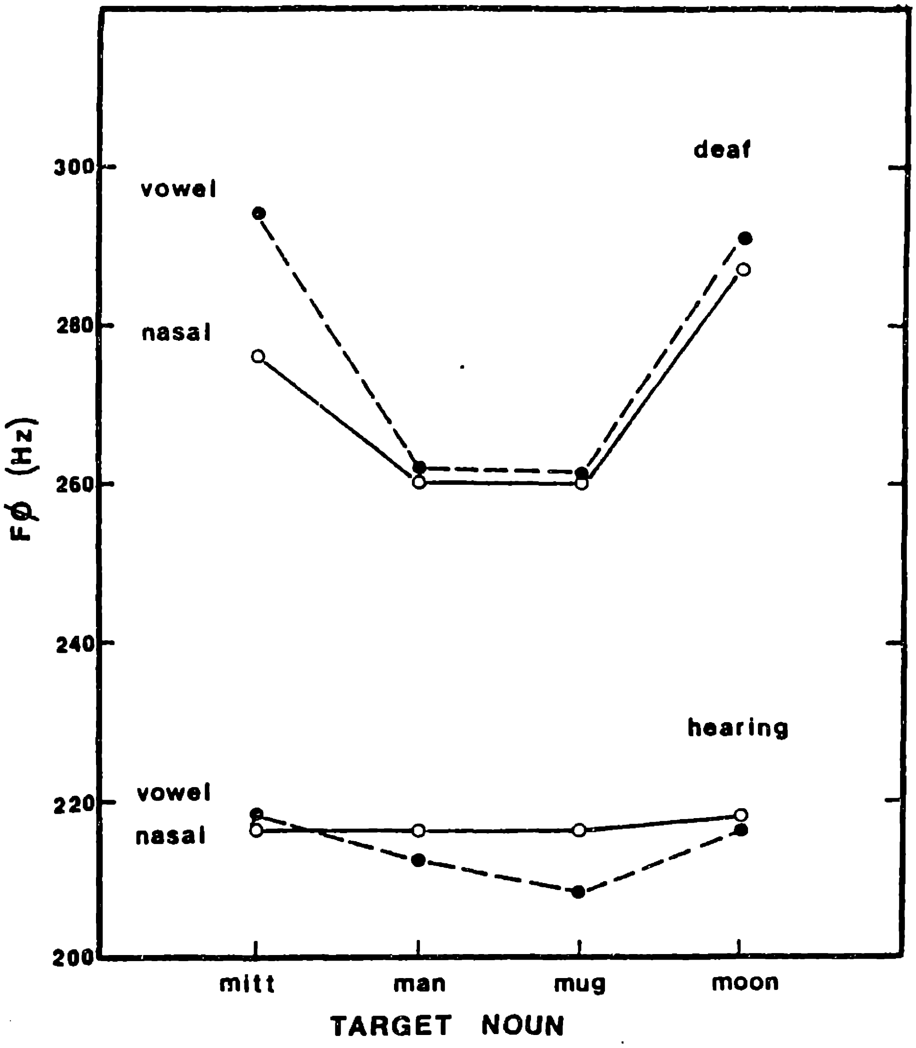


Figure 4.1 - F0 data for target nouns beginning with the nasal consonant /m/. F0 was measured at the "center" of the nasal consonant and at the "center" of the target vowel, as described in Section 2.4.1.

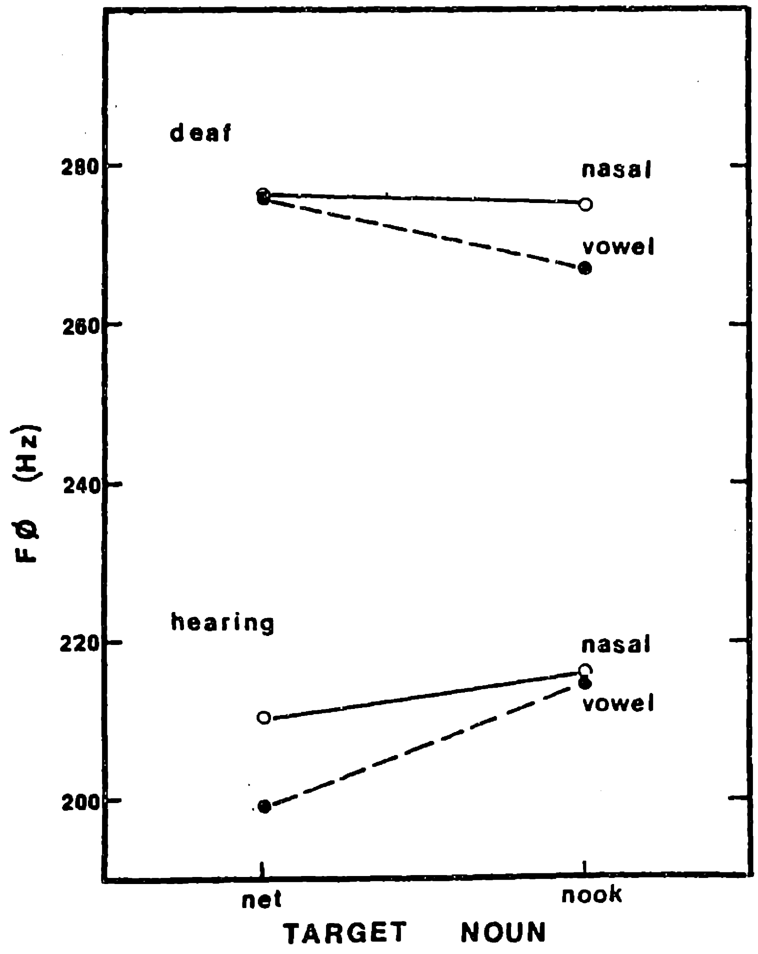


Figure 4.2 - F0 data for target nouns beginning with the nasal consonant /n/. F0 was measured at the "center" of the nasal consonant and at the "center" of the target vowel, as described in Section 2.4.1.

respect to place of articulation, eighteen of twenty /m/'s (90.0%), sixty-nine of eighty /n/'s (86.2%) and seven of twenty /ŋ/'s (35.0%) were accepted. Of these, one /ŋ/ token was ambiguous (i.e., it was identified as a nasal by one transcriber and as the velar stop /g/ by the other).

Figure 4.3 shows F0 data for the target nouns ending with nasal consonants. With the exception of the noun "time" as spoken by the hearing boys and girls, F0 tends to be lower at the center of the nasal consonant than at the center of the preceding vowel for both the deaf and hearing groups of speakers. As for the consonant /m/ in syllable-initial position, F0 on the nasal consonants appears to depend closely on F0 at the center of the adjacent (target) vowel for the deaf boys and girls. For the hearing speakers, variations in F0 on the nasal consonants, as a function of the target vowel, are somewhat less systematic.

A number of different F0 contours were imposed on the test phrases containing nouns with syllable-initial nasal consonants by both the deaf and hearing speakers. For the hearing group, the most common contours included: 1) relatively level or falling F0 throughout the nasal consonant and the target vowel; 2) a peak in F0 on the nasal consonant followed by falling F0 on the target vowel; and 3) a dip in F0 on the nasal consonant with a peak in F0 near the beginning of the following (target) vowel. Examples of the latter two

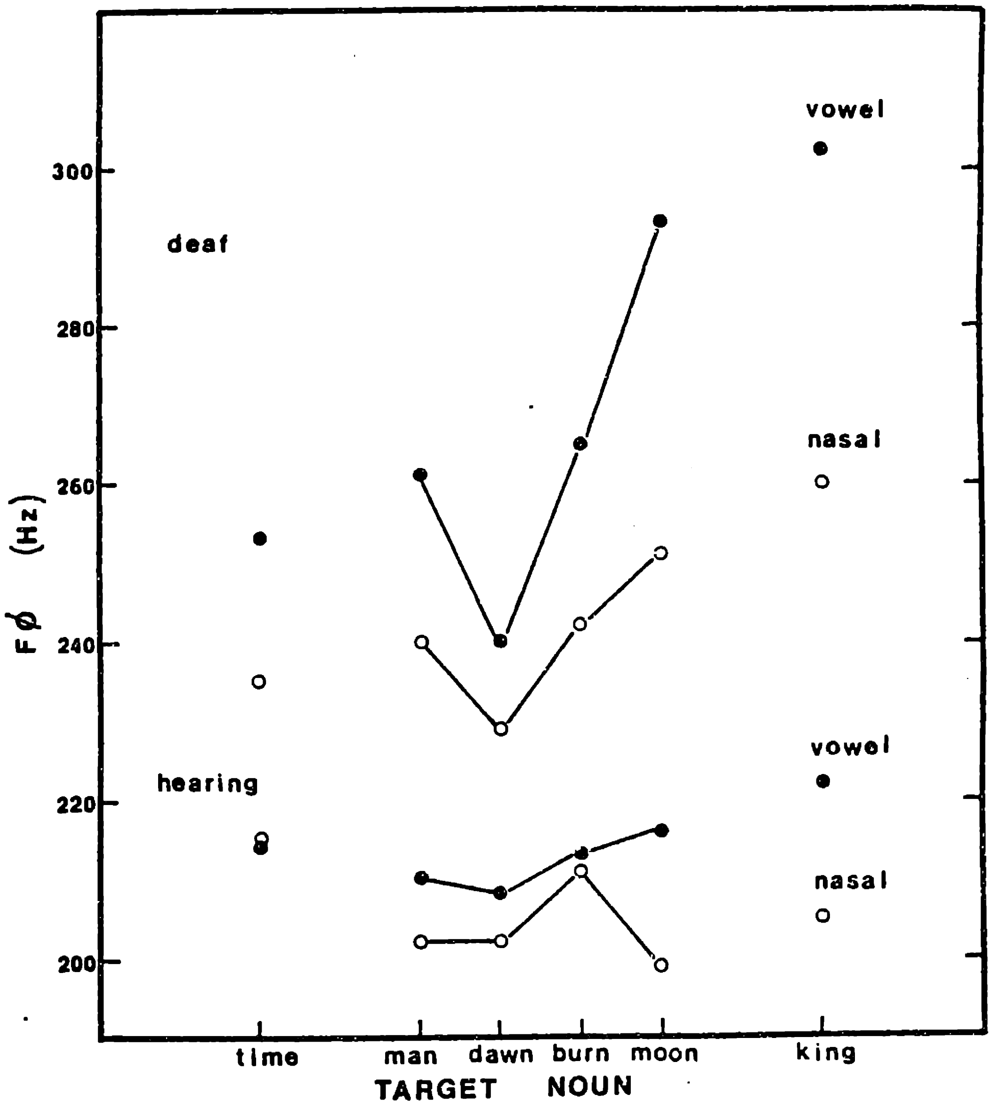


Figure 4.3 - F0 data for target nouns ending with the nasal consonants /m,n,ŋ/. F0 was measured at the "center" of the nasal consonant and at the "center" of the target vowel, as described in Section 2.4.1.

contours are shown in Figure 4.4.

For the most part, the F0 contours produced by the deaf boys and girls were similar in shape to those produced by the hearing speakers. In many instances, however, the amount of F0 change which occurred over the course of a syllable was considerably greater for the deaf speakers, and peaks in F0 tended to be located near the center (rather than near the beginning) of the target vowels. These similarities and differences can be seen by comparing the F0 contours of Figure 4.4 with those shown for two deaf speakers in Figure 4.5.

The deaf boys and girls also produced a number of "atypical" F0 contours (i.e., contours which were not observed, or observed only rarely, for the hearing speakers). As illustrated in the examples of Figure 4.6, these "atypical" contours included: 1) a relatively sharp rise in F0 from the nasal to the following (target) vowel; 2) a break in voicing at the consonant-vowel boundary, usually followed by higher F0 at the onset of the target vowel; and 3) a "glitch" in F0 at the consonant-vowel boundary, often with distinctive peaks (or plateaus) in F0 on both the nasal consonant and the target vowel.

Examination of wideband sound spectrograms indicated that the relatively abrupt (upward) changes in F0 associated with these "atypical" contours were not consistently related to

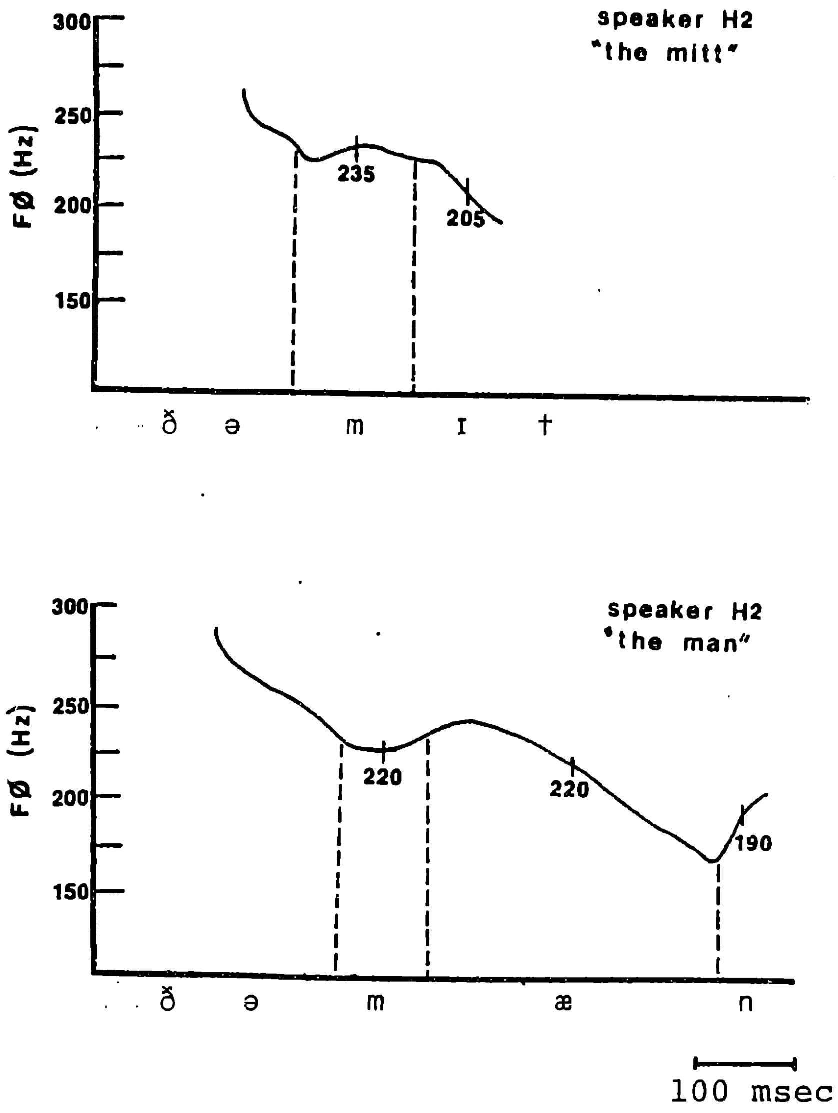


Figure 4.4 - FØ contours typical of those produced by the hearing speakers for target nouns beginning with nasal consonants. (See text for a description.)

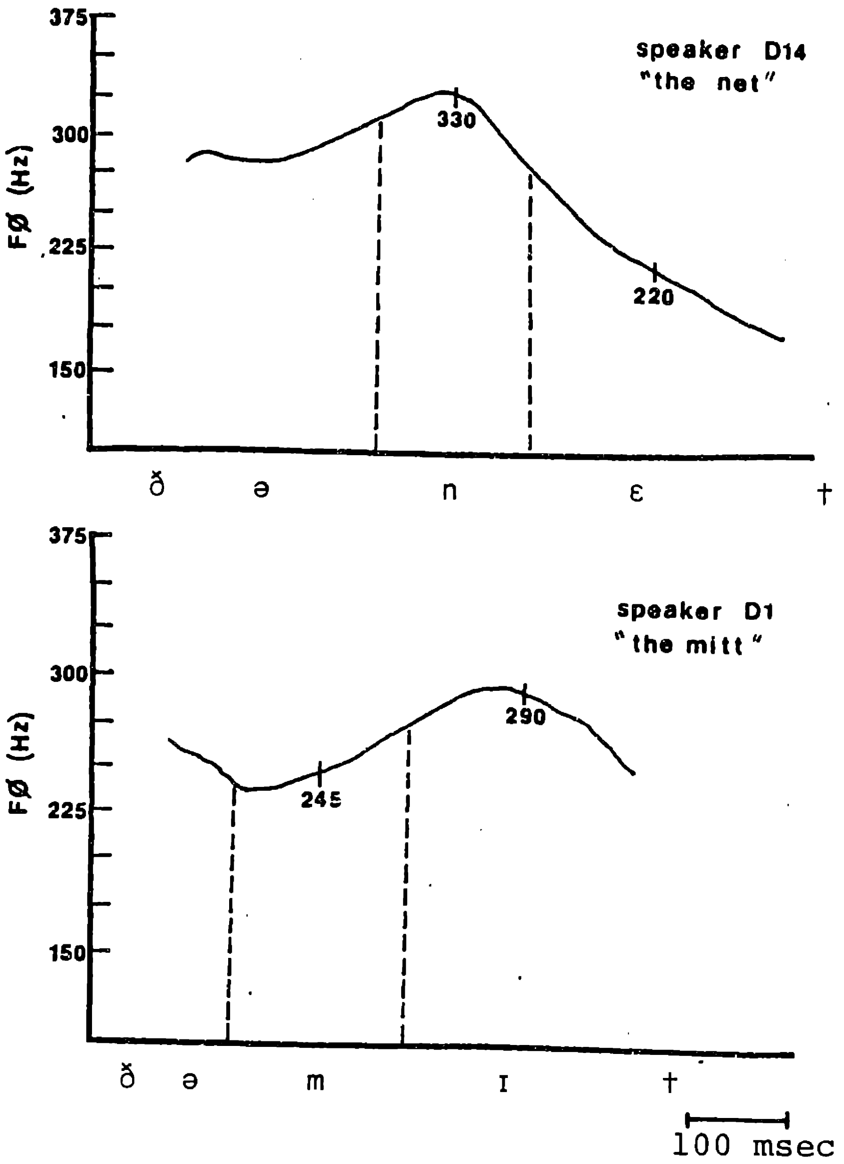


Figure 4.5 - F0 contours typical of those produced by the deaf speakers for target nouns beginning with nasal consonants. (See text for a description.)

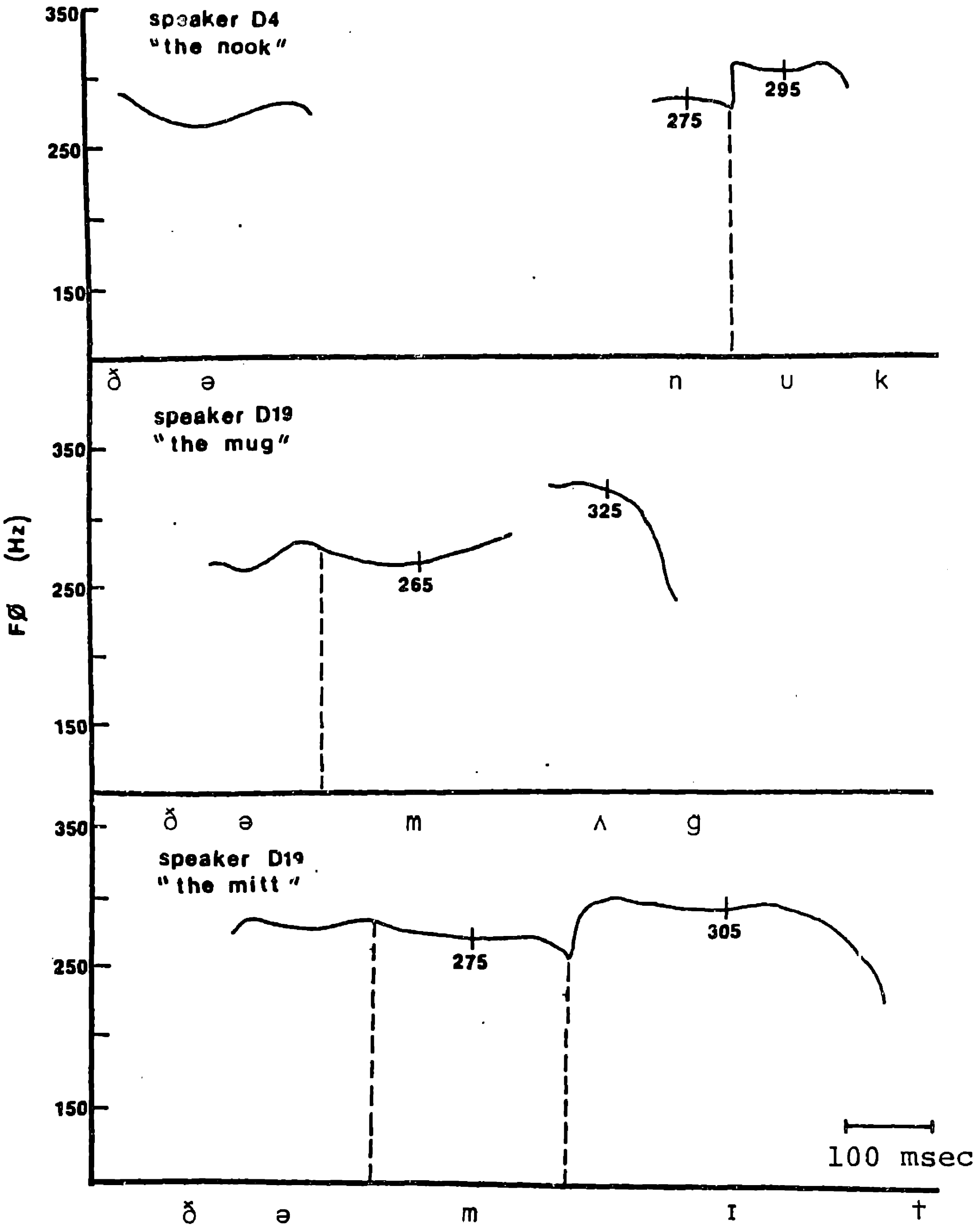


Figure 4.6 - Some "atypical" F0 contours produced by the deaf speakers for target nouns beginning with nasal consonants. (See text for a description.)

(downward) changes in the frequency of F1, as would be predicted by an acoustic-coupling hypothesis. In many instances, for example, F1 was low throughout the nasal consonant and the target vowel while, in others, both F0 and F1 increased in frequency near the consonant-vowel boundary. Only rarely did F1 appear to be lower (and F0 higher) on the target vowel than on the preceding consonantal segment.

In the case of nasals in syllable-final position, F0 could be rising, falling or level for both the deaf and hearing boys and girls. The deaf speakers, however, typically produced consonants which were much longer in duration than those produced by the hearing speakers and occasionally imposed a pronounced peak (or plateau) on the (lengthened) nasal segments. As indicated in the examples of Figures 4.7 and 4.8, the (upward) changes in F0 associated with such contours could be gradual or relatively abrupt or could be part of a "glitch" in the F0 contour near the vowel-consonant boundary.

Again, examination of wideband spectrograms showed no systematic relationship between upward changes in F0 on a syllable-final nasal consonant and changes in the frequency of F1. Increases in F0 were observed for nasal consonants following both low and high vowels (i.e., vowels with high and low F1) and could begin during the vowel, near the vowel-consonant boundary or well after the onset of the nasal

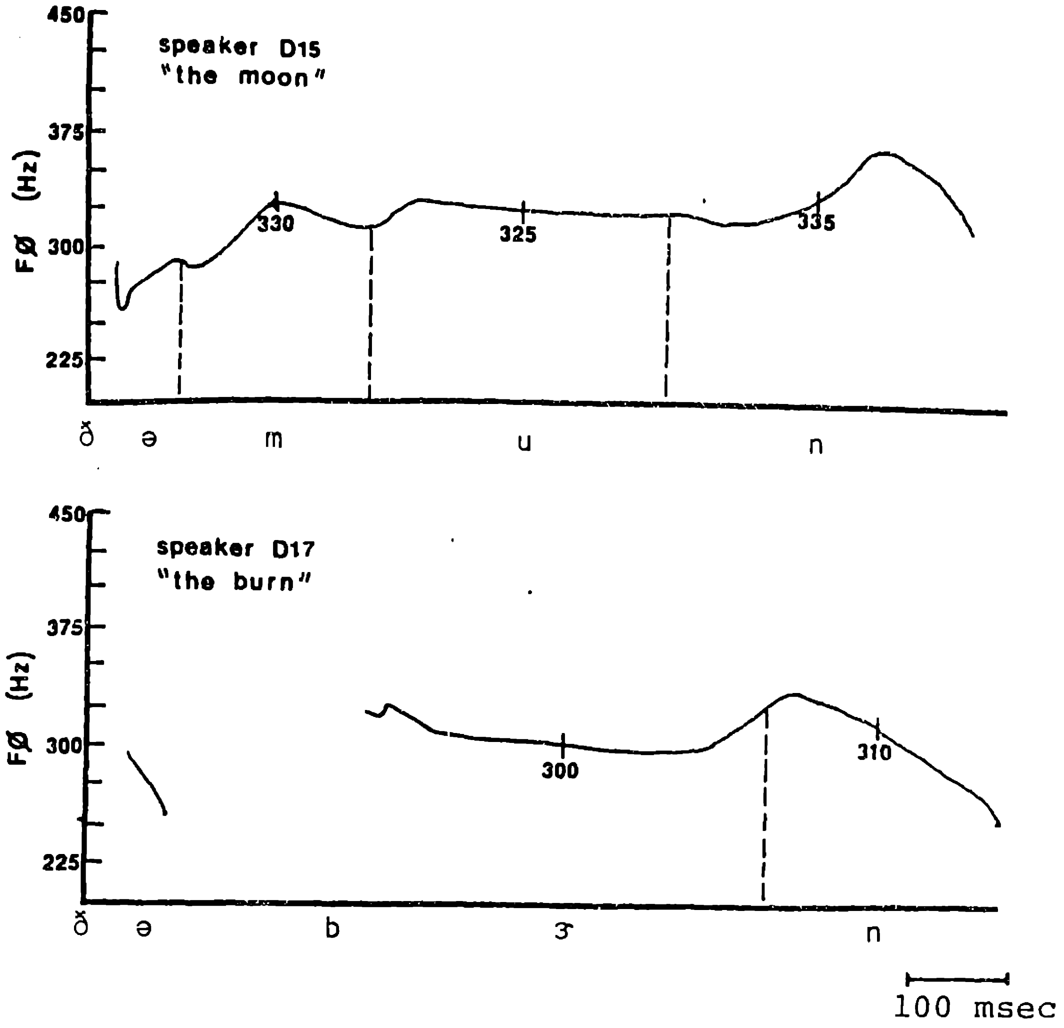


Figure 4.7 - Examples of FØ contours produced by the deaf speakers for target nouns ending with nasal consonants. (See text for a description.)

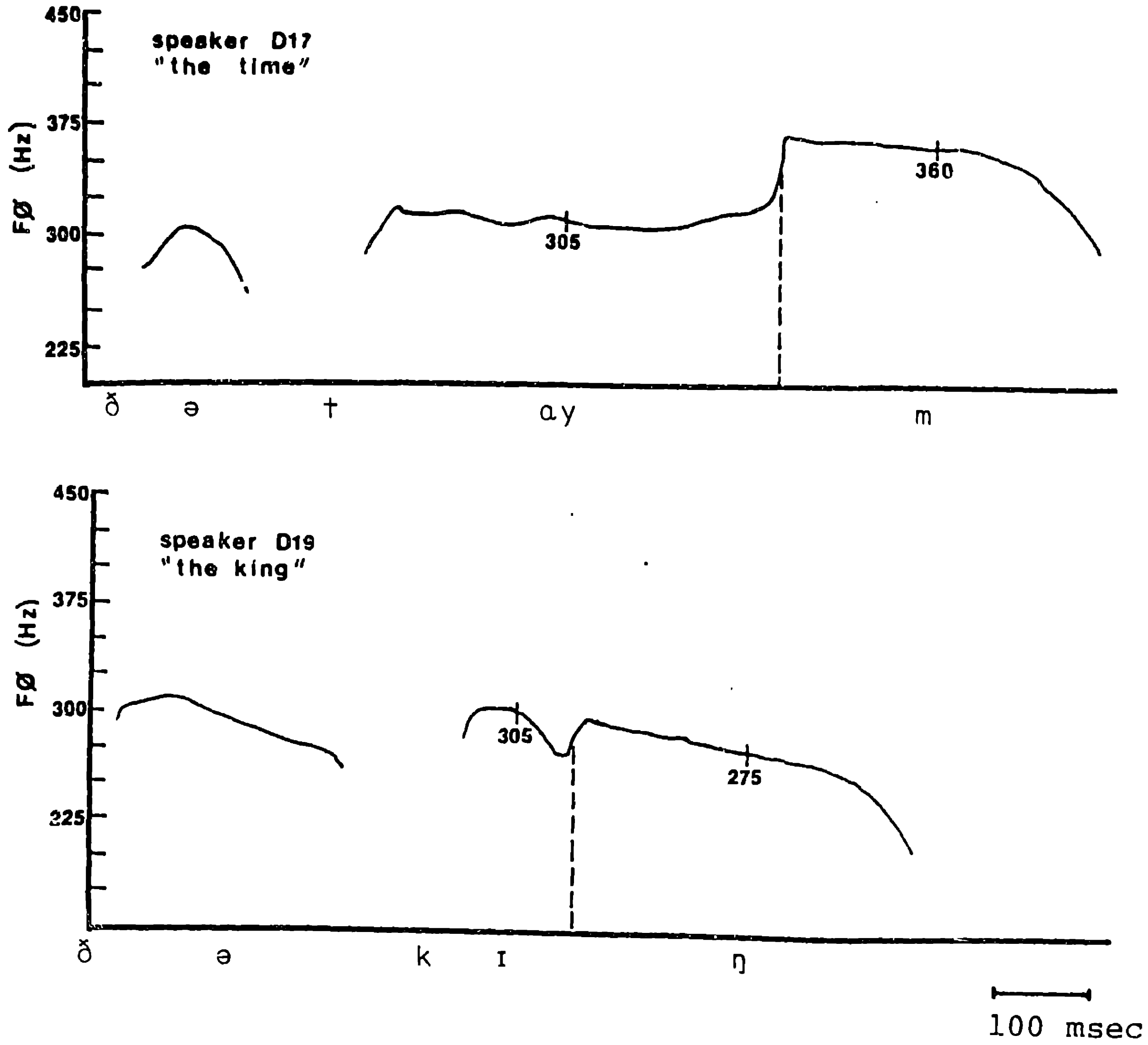


Figure 4.8 - Examples of FØ contours produced by the deaf speakers for target nouns ending with nasal consonants. (See text for a description.)

segment. These various situations are illustrated in the examples of Figures 4.7 and 4.8.

In order to obtain some information about the differences and/or similarities among the groups of deaf speakers described in the preceding chapter (Section 3.2.2), F0 data for each group was averaged for target nouns containing the high vowels /i/ and /u/ (preceded by /m/) and for nouns containing the low vowels /æ/ and /^/ (preceded by /m/). These averaged data, together with comparable data for the hearing controls, are shown graphically in Figure 4.9. (Numerical data as well as the number of tokens on which each averaged F0 measurement is based are provided in Table A2.3 in Appendix 2.)

For the most part, the differences in the amount of F0 variability and the F0 level used by the speakers in Groups A-1, A-2, B and C are consistent with those observed in the preceding chapter. In particular, the difference between F0 on target nouns with high vowels and target nouns with low vowels is greatest for the speakers of Groups A-2 and C (on consonants as well as on vowels), and F0 is higher than that used by the hearing controls, even for the low vowels /æ/ and /^/. The speakers in Group A-1 again produce target nouns with both high and low vowels with very high F0, while the speakers in Group B are most like normal with respect to both F0 level and vowel-related variability in F0.

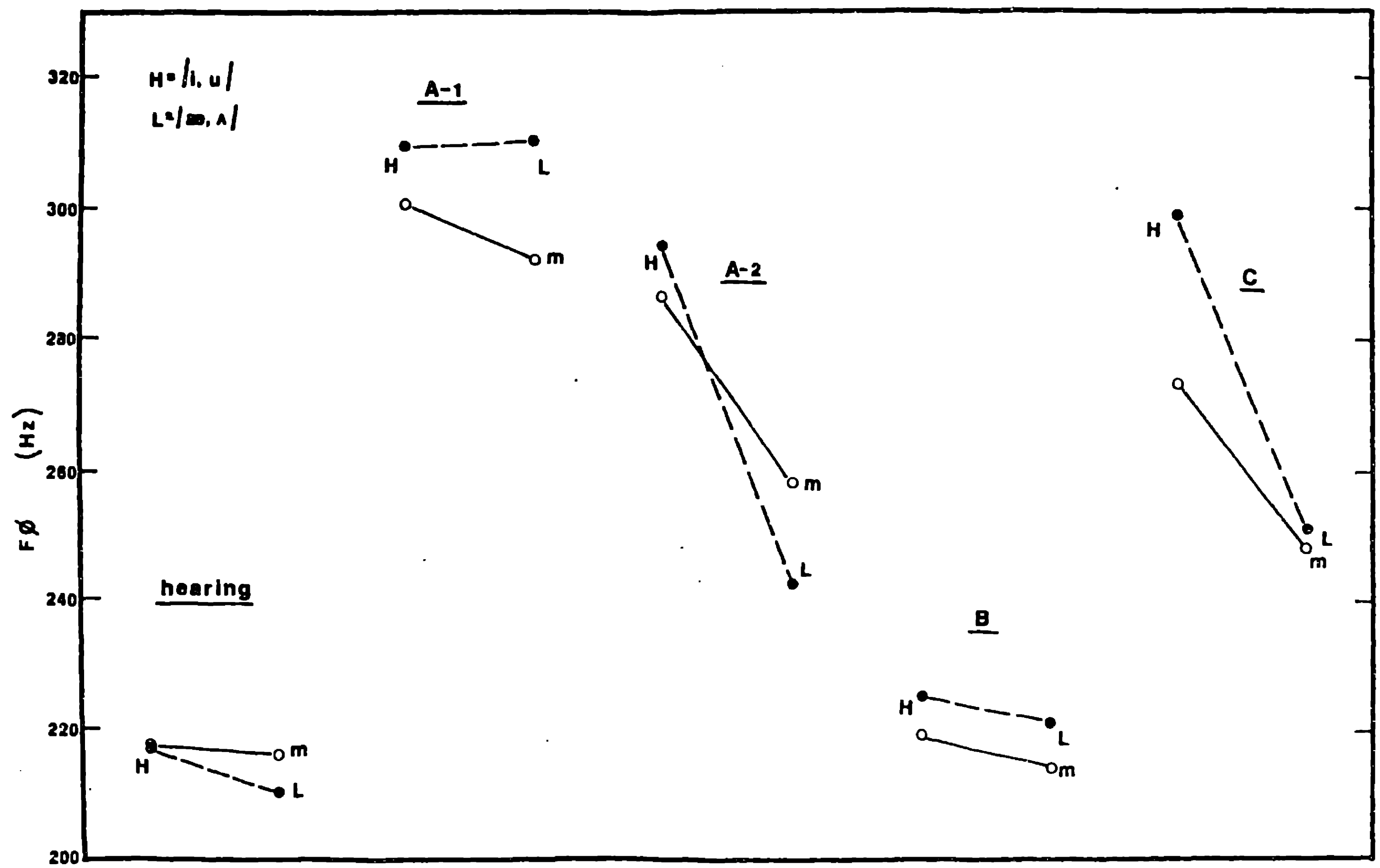


Figure 4.9 - Grouped F0 data for target nouns containing high (H) and low (L) vowels preceded by the nasal consonant /m/. F0 was measured at the "center" of the nasal consonant and at the "center" of the target vowel, as described in Section 2.4.1.

For the deaf speakers of Groups A-2, B and C, F0 on the nasal consonant /m/ again appears to depend closely on F0 on the following (target) vowel. F0 is somewhat lower on /m/'s preceding low vowels than on /m/'s preceding high vowels for the speakers of Group A-1, while F0 on the target vowels remains approximately constant. Except in the case of nouns containing low vowels, as produced by the speakers of Group A-2, F0 is lower on the /m/ than on the target vowel for the deaf speakers.

The F0 data shown in Figure 4.9 (and in Figures 4.1 through 4.3) suggest that, for the majority of deaf boys and girls, F0 on nasal consonants depends less on the presence of a low-frequency first formant than on the height of the adjacent target vowel. This observation, together with the observations that F0 tends to be lower on nasal consnants than on the target vowels (in most instances) and that abrupt (upward) changes in F0 are not consistently related to (downward) changes in the frequency of F1, would seem to argue <u>against</u> an acoustic-coupling coupling mechanism for vowel-related F0 change (assumption 1, Section 4.1.1).

Before the F0 data for the remaining consonantal contexts are described, two comments can be made about the "atypical" F0 contours produced by the deaf boys and girls. First, examination of the phonetic transcriptions showed that a stop consonant was often inserted between the nasal consonant and the target vowel in utterances for which an abrupt change in F0 occurred in the vicinity of a consonant-vowel boundary (e.g., Figure 4.6, top). This observation suggests the possibility that, at least in some instances, such changes in F0 might be associated with a build-up of oral air pressure during stop production [Ste-pc].

Second, it is possible that those contours characterized by breaks or glitches in F0 near a consonant-vowel or vowel-consonant boundary might reflect an improper concatenation of adjacent speech sounds. For example, the deaf speakers might, in some cases, tend to produce the nasal consonant and the target vowel as distinct segments, each with a characteristic F0 contour (e.g., a peak or plateau in F0 on each segment, as in the lower contours of Figures 4.6 and 4.8) or with a pause and, thus, a break in voicing (as in the center contour of Figure 4.6) between the two speech sounds. Similar problems with word and syllable concatenation have been reported for deaf speakers by Bernstein, Rollins and Stevens [BerRol78].

4.1.3 /r,l,w/

Nouns containing the sonorants /r,l,w/ (Category 4 in Table 2.3) were originally included in the test corpus in order to obtain additional information about the influence of a low-frequency first formant on the deaf speakers control of F0. Preliminary acoustic measurements had suggested that some deaf speakers tended to produce nasalized versions of these consonants (as well as nasal consonants and high vowels) with inappropriately high F0, and the hypothesis was made that this high F0 might be the result of an acoustic-coupling effect of F1 on F0 (Section 1.2).

An analysis of the F0 contours produced by the deaf speakers for target nouns containing /r,l,w/ did not support these preliminary observations. In most instances, F0 on both syllable-initial /r,l,w/ and syllable-final /l/ was lower than F0 on the adjacent target vowel, whether or not the consonant was transcribed as a nasal (or nasalized) segment. Furthermore, as for nouns containing nasal consonants, changes in F0 in the vicinity of a consonant-vowel (or vowel-consonant) boundary were not consistently related to changes in the frequency of F1.

Examination of the phonetic transcriptions for this category of target nouns indicated that the consonants /r/ and /l/ were usually produced incorrectly by the deaf boys and girls. Only fourteen of sixty /r/'s, twenty-two of sixty syllable-initial /l/'s and twenty-two of sixty syllable-final /l/'s were labelled as such, and often by only one of the two transcribers. The sonorant /w/ was transcribed as such (by at least one transcriber) in fifty-six of eighty cases.

Because so few of the sonorants (particularly /r/ and /l/) were produced correctly, it was difficult to define criteria for accepting a token based on adequacy of consonant articulation. Therefore, a decision was made to measure F0 in all cases in which voiced consonantal segments could be identified on the appropriate sound spectrograms. One hundred and forty-seven of two hundred (73.5%) syllable-initial tokens and fifty-one of sixty (85.0%) syllable-final tokens met this arbitrary criterion. (See Tables A2.4 through A2.6 in Appendix 2 for the number of tokens accepted for each target noun.)

F0 data for target nouns beginning with /r,l,w/ and ending with /l/ are shown in Figures 4.10 and 4.11 respectively. (F0 measurements were again made at the "centers" of the consonantal segments and at the "centers" of the target vowels, as described in Section 2.4.1.) For the deaf group of speakers, F0 on both syllable-initial and syllable-final consonants appears to depend closely on F0 at the center of the adjacent (target) vowel. With the exception of the /r/ in "rug", F0 on the consonant is comparable to or lower than F0 on the vowel--much lower for some syllable-initial consonants followed by high vowels (e.g., the /l/ in "leak" or the /r/ in "root") and in the case of syllable-final /l/.

For the hearing group of speakers, variations in F0 on syllable-initial /r,l,w/ (as a function of the target noun) are small and appear to be relatively independent of F0 on the following vowel (Figure 4.10). F0 at the center of these consonants is higher than or comparable to F0 at the center of the vowel, except in the case of nouns containing the high vowels /i/ and /u/ (i.e., "leak" and "root"). Cross-vowel changes in F0 on syllable-final /l/ are similar in direction to, but somewhat greater in magnitude than, those on the preceding target vowels (Figure 4.11), and F0 on the /l/ is lower than, higher than and comparable to F0 on the vowel in the nouns "well", "wall" and "wool" respectively. It should be noted, however, that syllable-final /l/'s were often hard to locate on the sound spectrograms for the hearing speakers, and these latter data are based on a relatively small number of measurements (see Table A2.5).

The various types of F0 contours (typical and atypical) described for nouns with nasal consonants in syllable-initial position were also observed for nouns beginning with /r,l,w/ by the deaf and hearing boys and girls (see Figures 4.4 through 4.6). Again, there was no systematic relationship

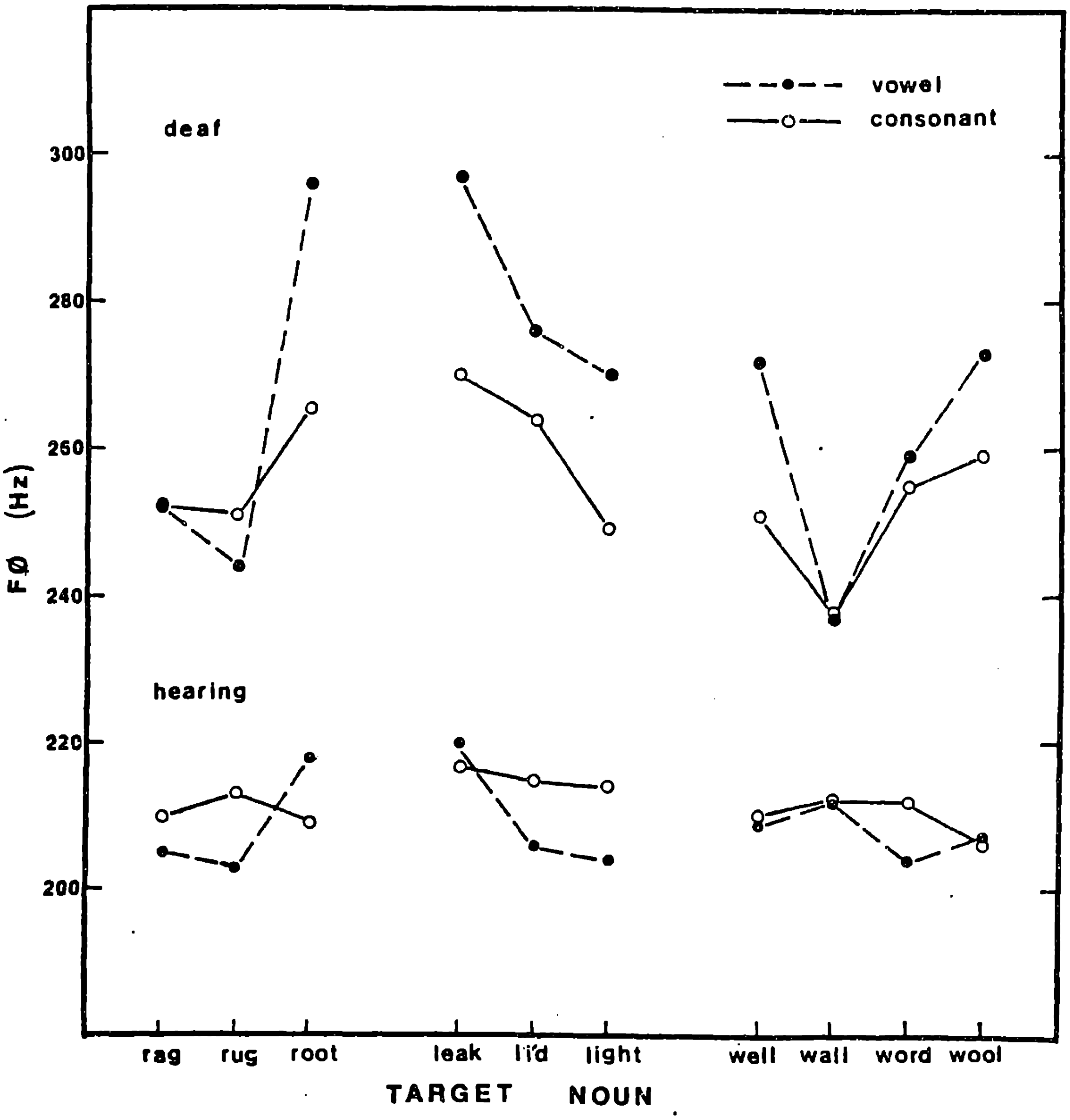


Figure 4.10 - FØ data for target nouns beginning with the sonorants /r,l,w/. FØ was measured at the "center" of the consonant and at the "center" of the target vowel, as described in Section 2.4.1.

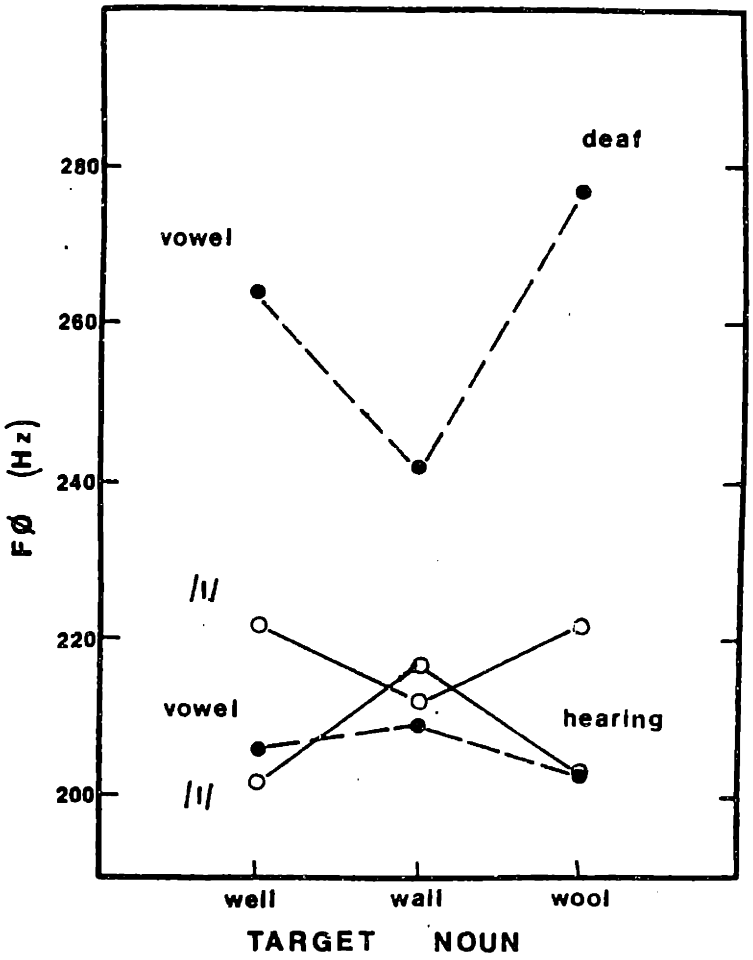


Figure 4.11 - FØ data for target nouns ending with the sonorant /l/. FØ was measured at the "center" of the /l/ and at the "center" of the target vowel, as described in Section 2.4.1.

between (upward) changes or breaks in F0 in the vicinity of consonant-vowel boundaries and changes in the frequency of F1. F0 was almost always falling on syllable-final /l/ for the deaf boys and girls, and could be rising, falling or level for the hearing speakers.

Figure 4.12 shows F0 data for the speakers of Groups A-1, A-2, B, C and for the hearing controls for target nouns beginning with /r,l,w/. As in the case of syllable-initial /m/ (Section 4.1.2), these data have been averaged for target nouns containing high (/i,I,u/) and low (/æ ,ɑ ,ʌ /) vowels. (The corresponding numerical data and numbers of tokens are provided in Table A2.7 in Appendix 2.) Examination of Figure 4.12 indicates that, for these particular consonantal contexts, all four groups of deaf speakers use an amount of vowel-related F0 variability that is greater than that used by the hearing controls. Again, the F0 level is highest for the speakers of Group A-1 and lowest (i.e., closest to normal) for the speakers of Group B.

In most instances, F0 on the consonantal segments is lower than F0 on the following target vowels for the deaf speakers. (Two exceptions are in nouns containing low vowels as produced by the speakers of Groups A-2 and C, for which F0 on the consonant is higher than and comparable to F0 on the vowel respectively.) For all four groups of deaf speakers, F0 on the consonant appears to depend on F0 on the following

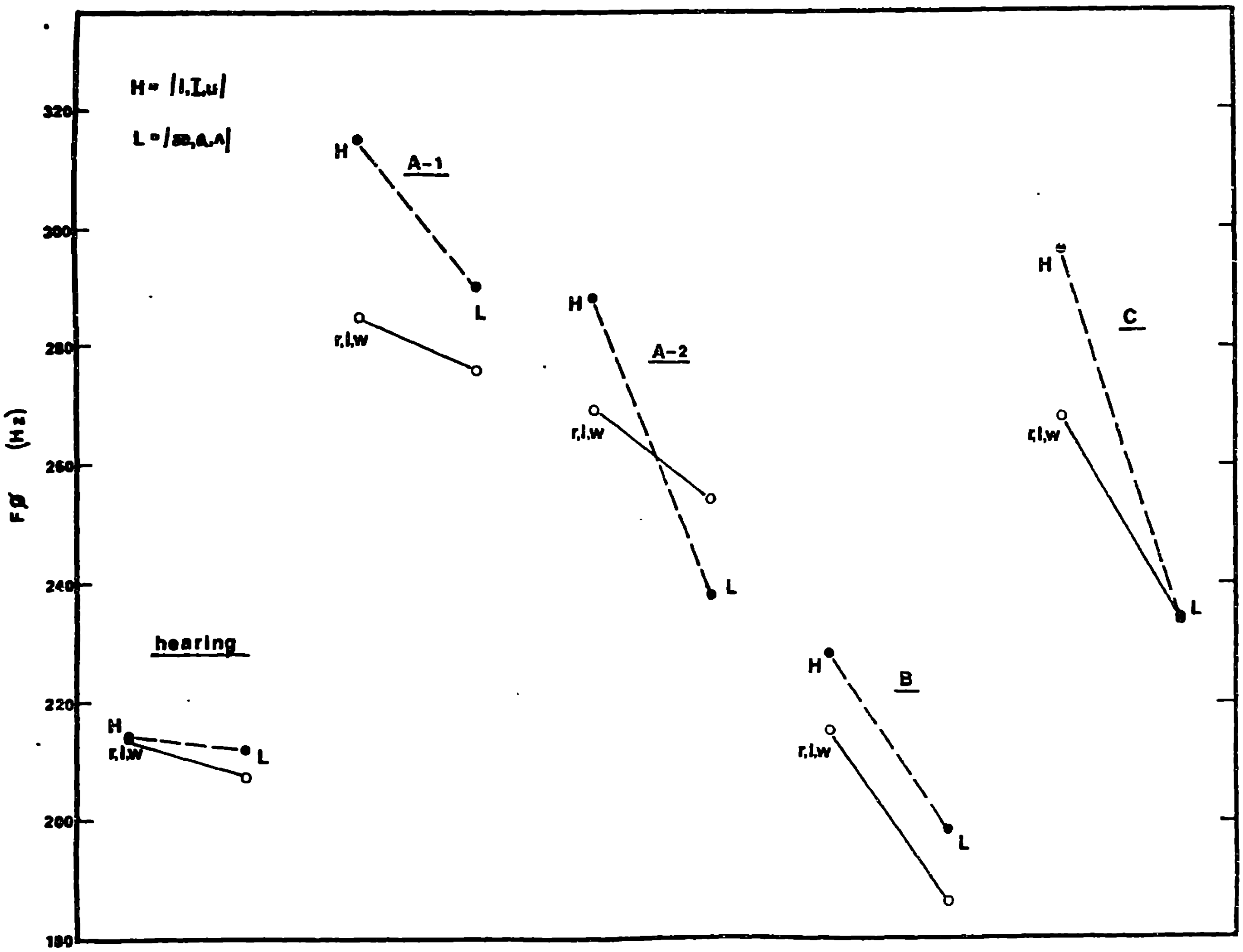


Figure 4.12 - Grouped F0 data for target nouns containing high (H) and low (L) vowels preceded by the sonorants /r,l,w/. F0 was measured at the "center" of the consonant and at the "center" of the target vowel, as described in Section 2.4.1.

(target) vowel. These observations, together with the lack of a systematic relationship between abrupt (upward) changes in F0 and (downward) changes in F1, would again seem to argue against an acoustic-coupling mechanism for vowel-related F0 change (assumption 1, Section 4.1.1).

4.1.4 Voiced Labial Stops

In order for a token of the labial stop /b/ to be accepted (for the deaf speakers), both transcribers had to agree on place and manner of consonant articulation and at least one of the transribers had to have labelled the stop as "voiced". Of the two hundred tokens possible, one hundred and forty-one (70.5%) met these criteria and, of the tokens accepted, fifty-six were "ambiguous" with respect to the voicing feature. (The number of tokens accepted for each target noun is provided in Tables A2.8 and A2.9 in Appendix A2.)

Figure 4.13 shows F0 data for target nouns beginning with the labial consonant /b/ for the deaf and hearing groups of speakers. (These data are based on F0 measurements made at points approximately ten and twenty milliseconds following vowel "onset" and at the "center" of the target vowel, as described in Section 2.4.1.) For both groups of speakers, F0 near the onset of the syllable appears to depend closely on F0 at the center of the following vowel. While the deaf speakers, on average, produce a wider range of cross-vowel F0 values, F0 tends to be higher for target syllables with high vowels than for target syllables with low vowels for both the deaf and hearing speakers.

The amount by which F0 changes over the course of the target nouns, on the other hand, is not consistently related to vowel height for either group of speakers. This can be seen more clearly in Figure 4.14, in which the average frequency difference between F0 measured 20 milliseconds after vowel onset (i.e., F0(20)) and F0 at the center of the vowel (i.e., F0(CV)) is plotted as a function of the vowel target. (The measure [F0(20) - F0(CV)] was computed for each speaker and then averaged.) Note, in particular, for the deaf boys and girls, that F0 changes by relatively small amounts over nouns containing both the high vowels /i/ and /u/ and the low vowel /ɑ/, although, for most speakers, changes in jaw position should be considerably greater in the latter vowel context.

Returning to Figure 4.13, one can see that, for the deaf speakers as a group, F0 tends to rise slightly near syllable onset (in most vowel contexts) and to drop to a lower value near the center of the target vowel. Examination of individual F0 contours, however, showed that F0 could be rising, falling or level near the start of a syllable and

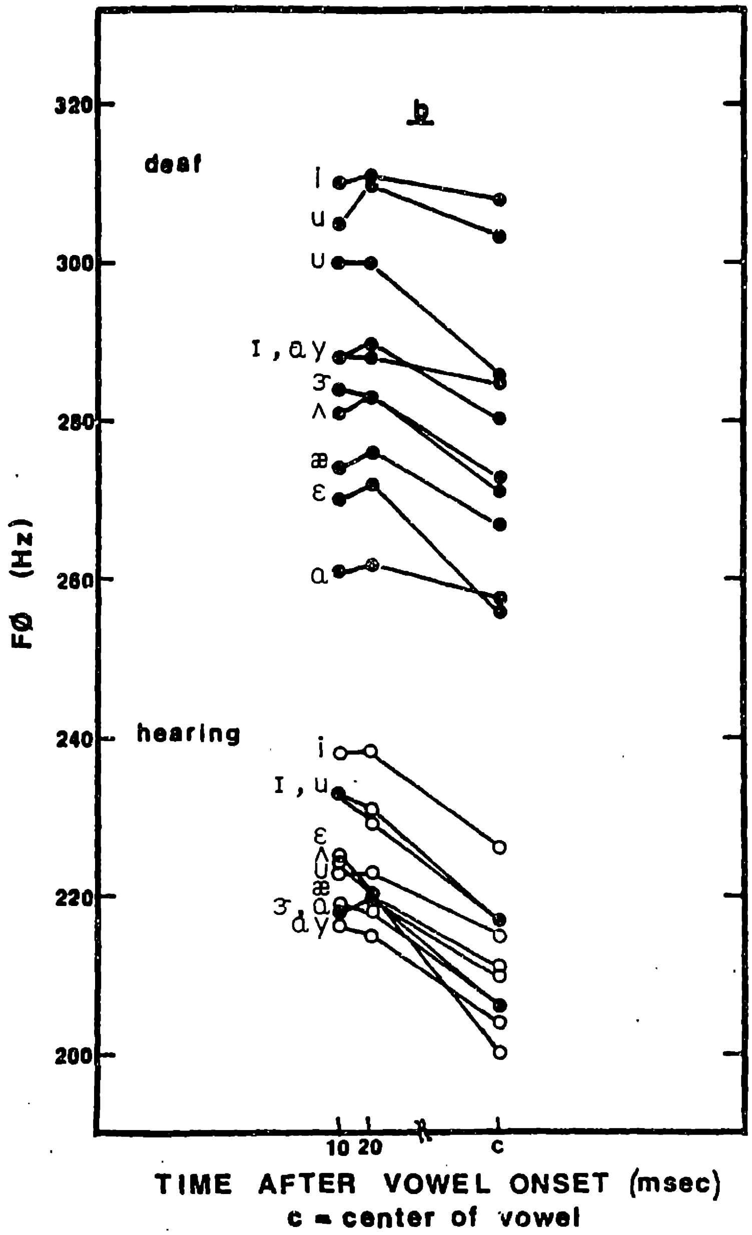


Figure 4.13 - F0 data for target nouns beginning with the labial stop consonant /b/.

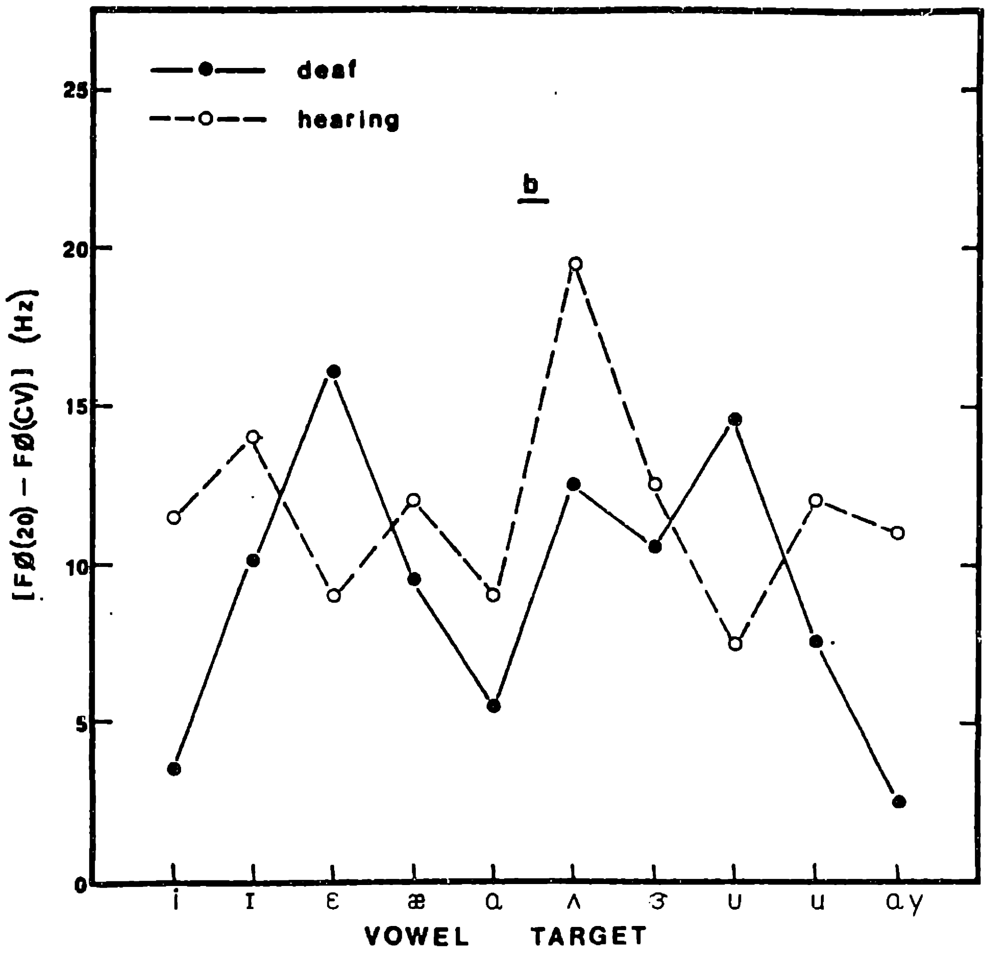


Figure 4.14 - [FØ(20) - FØ(CV)] as a function of the vowel target for nouns beginning with the labial stop consonant /b/. (See text for a definition of [FØ(20) - FØ(CV)].)

that, in a substantial number of cases (i.e., for approximately 30% of the acceptable tokens), F0 at the center of the vowel was actually higher than F0 near syllable onset. For the hearing speakers, F0 near syllable onset could again be rising, falling, or level, but F0 almost invariably fell to a lower value near the center of the target vowel.

It should be noted that, other things being equal, rising and falling F0 near syllable onset have most often been reported in the literature for syllables beginning with voiced and voiceless obstruents respectively [Lea73];[LehPet61]. However, the direction of F0 change following the release of a stop consonant can also depend on other factors, including stress (e.g., F0 will tend to rise for both voiced and voiceless stops when a syllable is stressed and to fall in both cases when the syllable is unstressed [Lea73]) and intonation. In the data just described, for example, the observation that F0 often tends to decrease over the entire course of target syllables beginning with the voiced labial /b/ (especially for the hearing speakers) is most likely a consequence of a falling intonation contour--in particular a phrase-final fall in F0--imposed on the target nouns.

Figure 4.15 shows F0 data, averaged for target nouns containing high (/i,I,u/) and low (/æ,ɑ,ʌ/) vowels, for the deaf speakers of Groups A-1, A-2, B and C and for the hearing controls. (The corresponding numerical data and numbers of tokens are provided in Table A2.10 in Appendix 2.) For the speakers of Group A-1, F0 is comparable (and high) for nouns containing both high and low vowels, and the amount of F0 change imposed on the syllables is small. For the remaining three groups of deaf speakers, the difference between F0 on high and low vowels is considerably greater than that used by the hearing controls, and F0 near syllable onset appears to be closely related to F0 at the center of the target vowel.

Figure 4.15 also shows that the amount of F0 change imposed on the target nouns with initial /b/ by a given group of speakers (either deaf or hearing) is comparable for high and low vowels, even though the amount of jaw movement involved in producing labial consonants followed by low vowels should be considerably greater. This observation, together with the vowel-dependent nature of F0 near syllable onset, suggests that jaw position is relatively <u>less</u> important than tongue height in determining the vowel-related variations in F0 produced by the deaf (and hearing) speakers (assumption 2, Section 4.1.1).

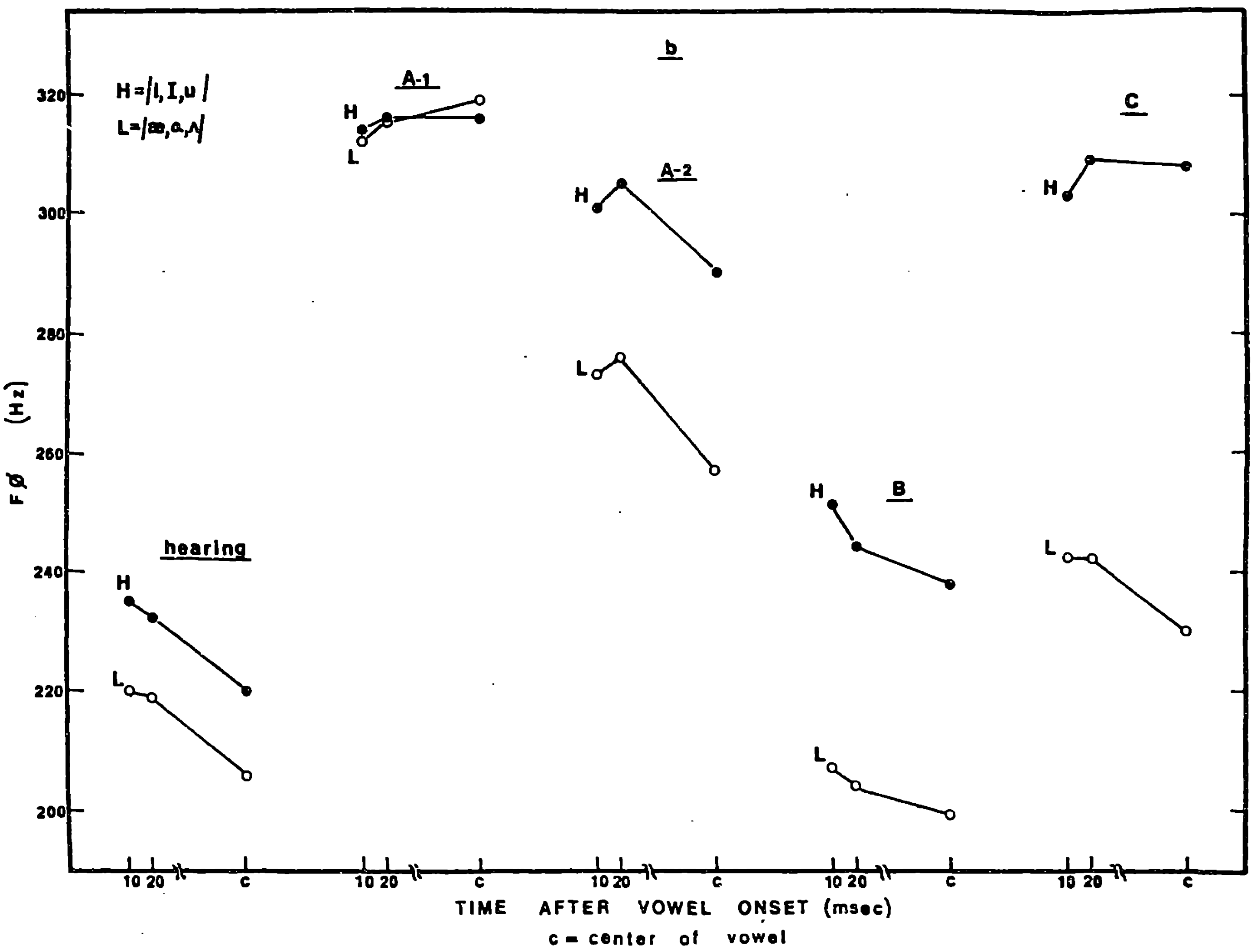


Figure 4.15 - Grouped F0 data for target nouns containing high (H) and low (L) vowels preceded by the labial stop consonant /b/.

4.1.5 Voiceless Stop Consonants

The criteria used in judging the acceptability of the stop consonants /p,t,k/ (as produced by the deaf speakers) were comparable to those used for the labial stop /b/: both transcribers had to agree on place and manner of consonant articulation and at least one transcriber had to have labelled the stop as "unvoiced". Of the two hundred and twenty tokens possible, one hundred and eighty (81.8%) met these criteria. With respect to place of articulation, sixty-nine of 80 /p/'s (86.2%), forty-six of sixty /t/'s (76.7%) and sixty-five of eighty /k/'s (81.2%) were accepted. The voicing feature was "ambiguous" for fourteen /p/, seven /t/ and ten /k/ tokens. (See Tables A2.11 and A2.12 in Appendix A2 for the number of tokens accepted for each target noun.)

Figure 4.16 shows F0 data for syllables beginning with voiceless stop consonants for the deaf and hearing groups of speakers. (Again, F0 was measured at points approximately ten and twenty milliseconds after vowel "onset" and at the "center" of the target vowel as described in Section 2.4.1.) As in the case of nouns beginning with the voiced labial stop /b/, F0 near the onset of the target syllables appears to depend closely on F0 at the center of the following vowel. Once more, F0 is higher for syllables containing high vowels than for syllables containing low vowels for both groups of

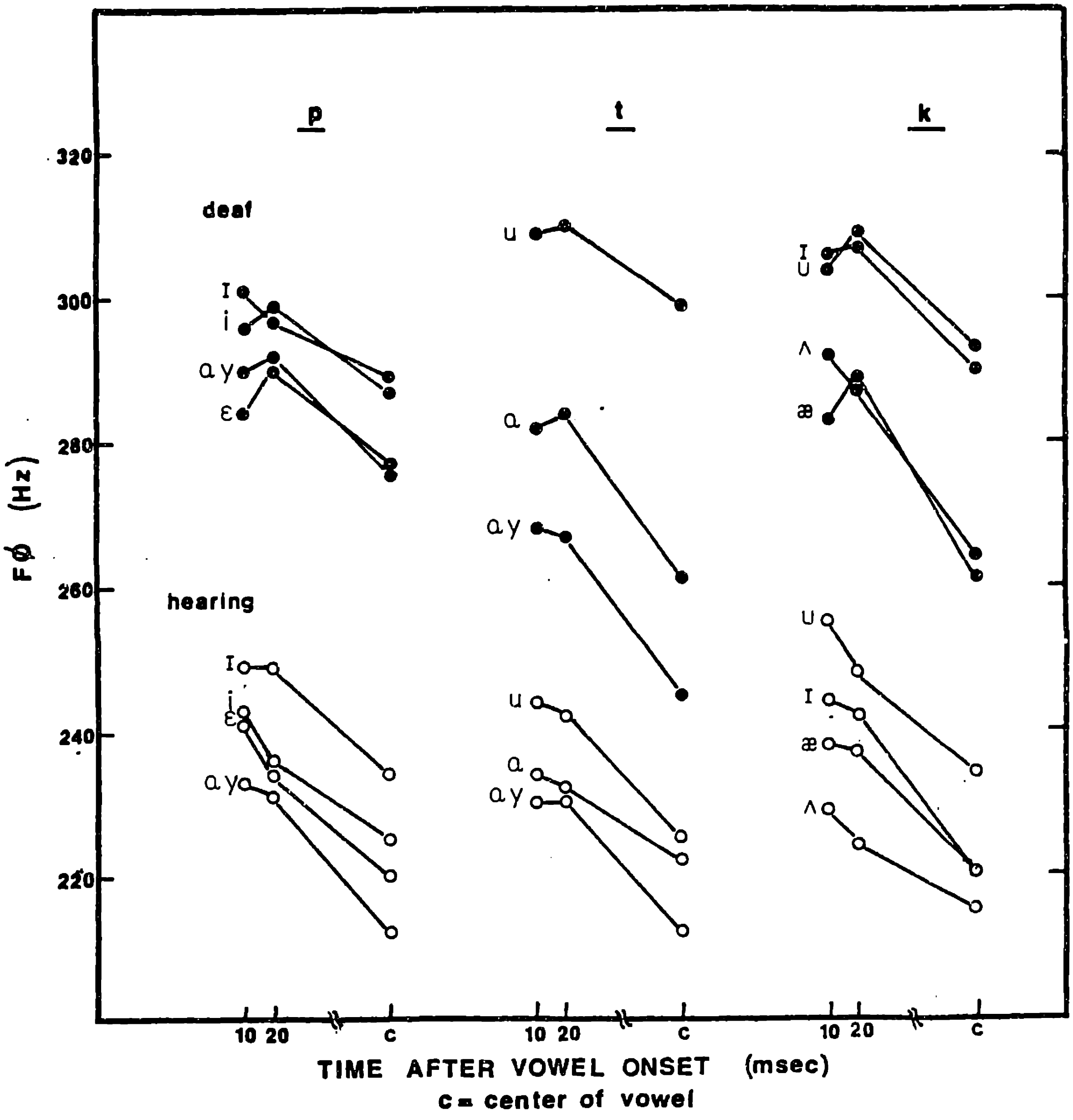


Figure 4.16 - FØ data for target nouns beginning with the voiceless stop consonants /p,t,k/.

speakers.

F0 data averaged across vowel contexts is shown for each place of consonant articulation in Figure 4.17. For both the deaf and hearing boys and girls, F0 near syllable onset is lowest for target nouns beginning with the alveolar consonant /t/. F0 is comparable near the onset of syllables beginning with the labial and velar consonants /p/ and /k/ for the hearing speakers; for the deaf speakers, F0 near syllable onset is somewhat higher for nouns beginning with the velar consonant /k/.

Given the apparent dependence of F0 near syllable onset on F0 on the following vowel, however, it is perhaps more appropriate to consider the relative amount of F0 change which occurs over the course of the target syllables (rather than simply F0 near the start of the syllable) in assessing the influence of place of consonant articulation on the deaf (or hearing) speaker's control of F0. In the case of the averaged data shown in Figure 4.17, the amount by which F0 changes over the target nouns remains relatively constant across consonantal contexts for the hearing boys and girls. For example, the average frequency difference between F0 measured 20 milliseconds after vowel onset (i.e., F0(20)) and F0 at the center of the vowel (i.e., F0(CV)) is 15 Hz, 15 Hz and 16 Hz for nouns beginning with /p/, /t/ and /k/ respectively. For the deaf speakers, on the other hand, F0 changes by a

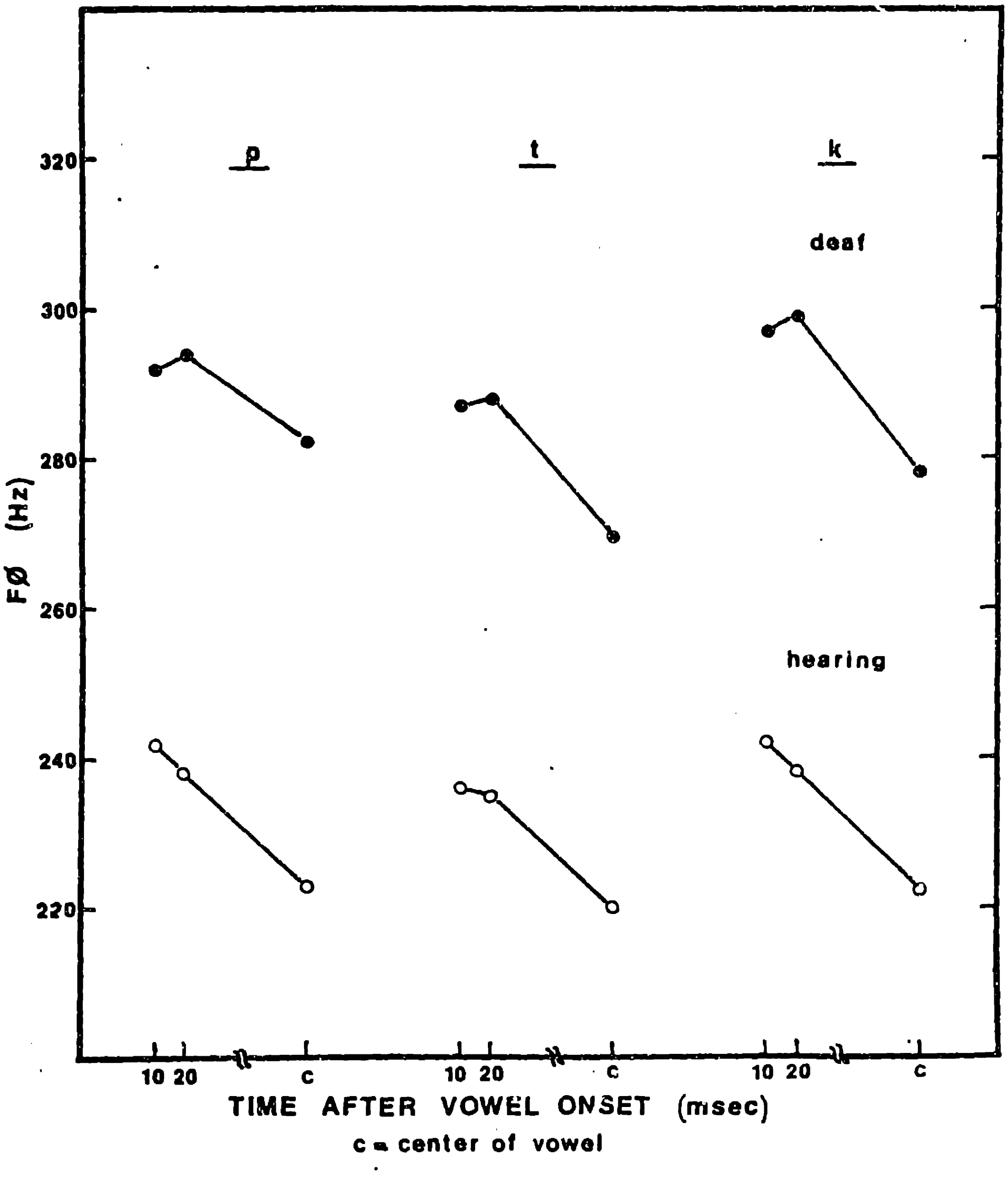


Figure 4.17 - Averaged FØ data for target nouns beginning with the voiceless stop consonants /p,t,k/.

greater amount, on average, over the course of nouns beginning with /t/ (19 Hz) and /k/ (21 Hz) than with /p/ (12 Hz).

In the preceding section, it was noted that the amount of F0 change imposed on *individual* target nouns beginning with the voiced labial stop /b/ was not consistently related to vowel height for either the deaf or hearing speakers (see Figure 4.14). Examination of Figure 4.18 indicates that the same is true for nouns beginning with the voiceless labial stop /p/ (for both groups), as well as for /t/ and for /k/ for the hearing speakers. (The measure plotted in this figure is again [F0(20) - F0(CV)] averaged across speakers.) For the deaf boys and girls, however, F0 changes by a much greater amount over the course of nouns containing low vowels (i.e., /ɑ,ɑy,æ,ʌ/) than nouns containing high vowels (i.e., /u,ʊ,I/) when the initial consonant is /t/ or /k/. These observations suggest that constraints on the position of the tongue may play an important role in controlling (i.e., increasing) F0 for at least some of the deaf speakers. (Note that for vowels preceded by the alveolar consonant /t/ the position of the tongue tip would normally be constrained near syllable onset, although there might be some constraint on the tongue body as well [Per69].)

The overall shapes of the F0 contours imposed on nouns beginning with voiceless stop consonants were, for the most part, similar to those observed for nouns beginning with the

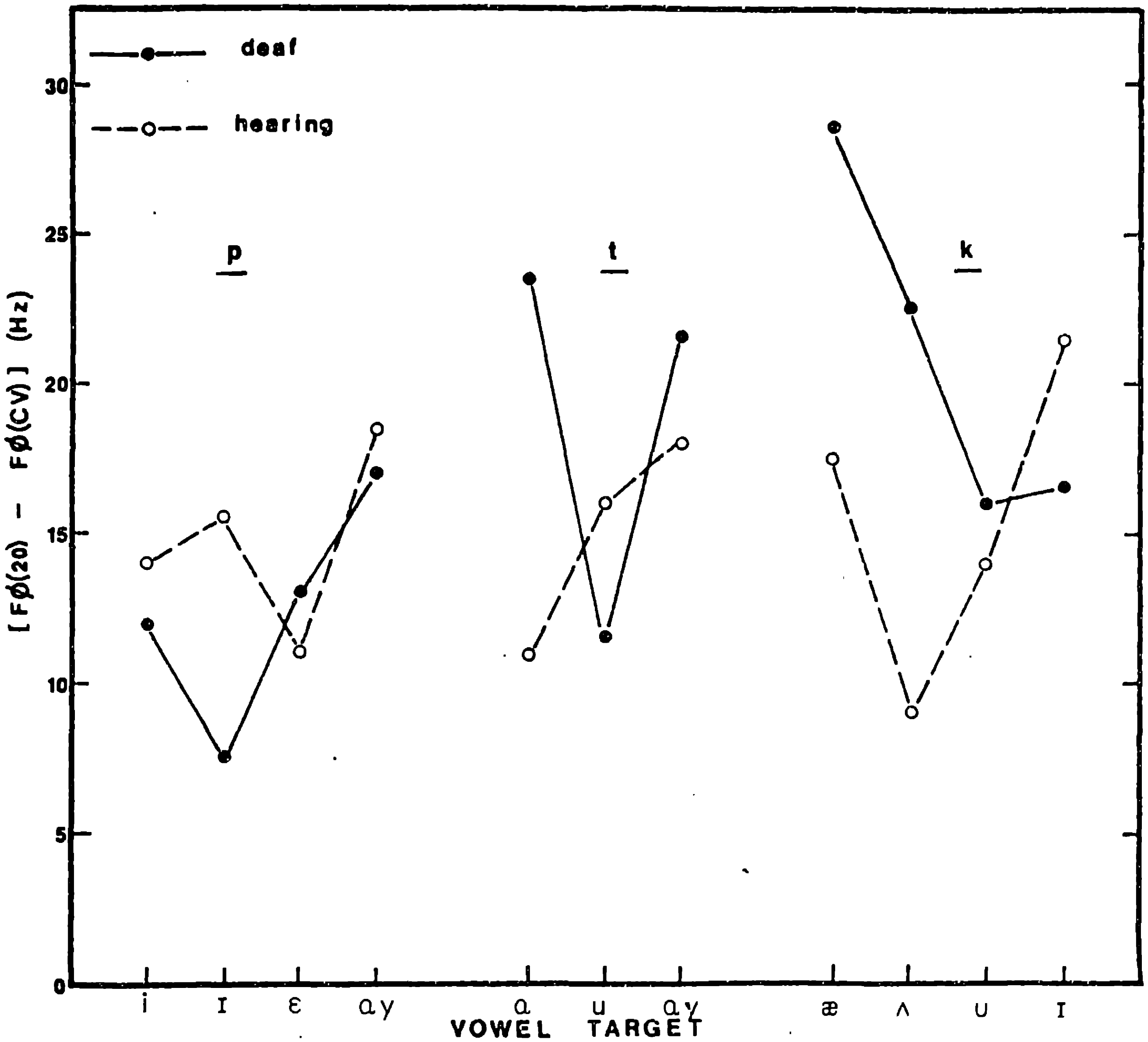


Figure 4.18 - [F0(20) - F0(CV)] as a function of the vowel target for nouns beginning with the voiceless stop consonants /p,t,k/. (See text for a definition of [F0(20 - F0(CV)].)

voiced labial stop /b/. As can be seen in Figures 4.16 and 4.17, F0 tends to rise slightly near syllable onset (for most vowel contexts) for the deaf speakers as a group and to drop to a lower value near the center of the target vowel. Again, however, examination of individual F0 contours showed that F0 could be rising, falling or level near the start of a syllable and that for approximately 25% of the acceptable tokens, F0 at the center of the target vowel was higher than F0 near syllable onset. For the hearing group of speakers, F0 could also be rising, falling or level near syllable onset and, again, F0 almost invariably fell to a lower value near the center of the target vowel. (See Section 4.1.3 for a discussion of factors which can contribute to the production of these various types of contours.)

Figure 4.19 compares F0 data, averaged for nouns containing high (/i,I,u/) and low (/æ,ɑ,ʌ/) vowels, for the deaf speakers of Groups A-1, A-2, B and C and for the hearing controls. (The corresponding numerical data and numbers of tokens are provided in Table A2.13 in Appendix 2). For all group of speakers (deaf and hearing), F0 near the onset of the target nouns again appears to depend upon F0 at the center of the vowel. The difference between F0 for nouns containing high and low vowels is relatively small for the speakers of Group A-1 and for the hearing controls, and relatively large for the remaining three groups of deaf speakers (particularly

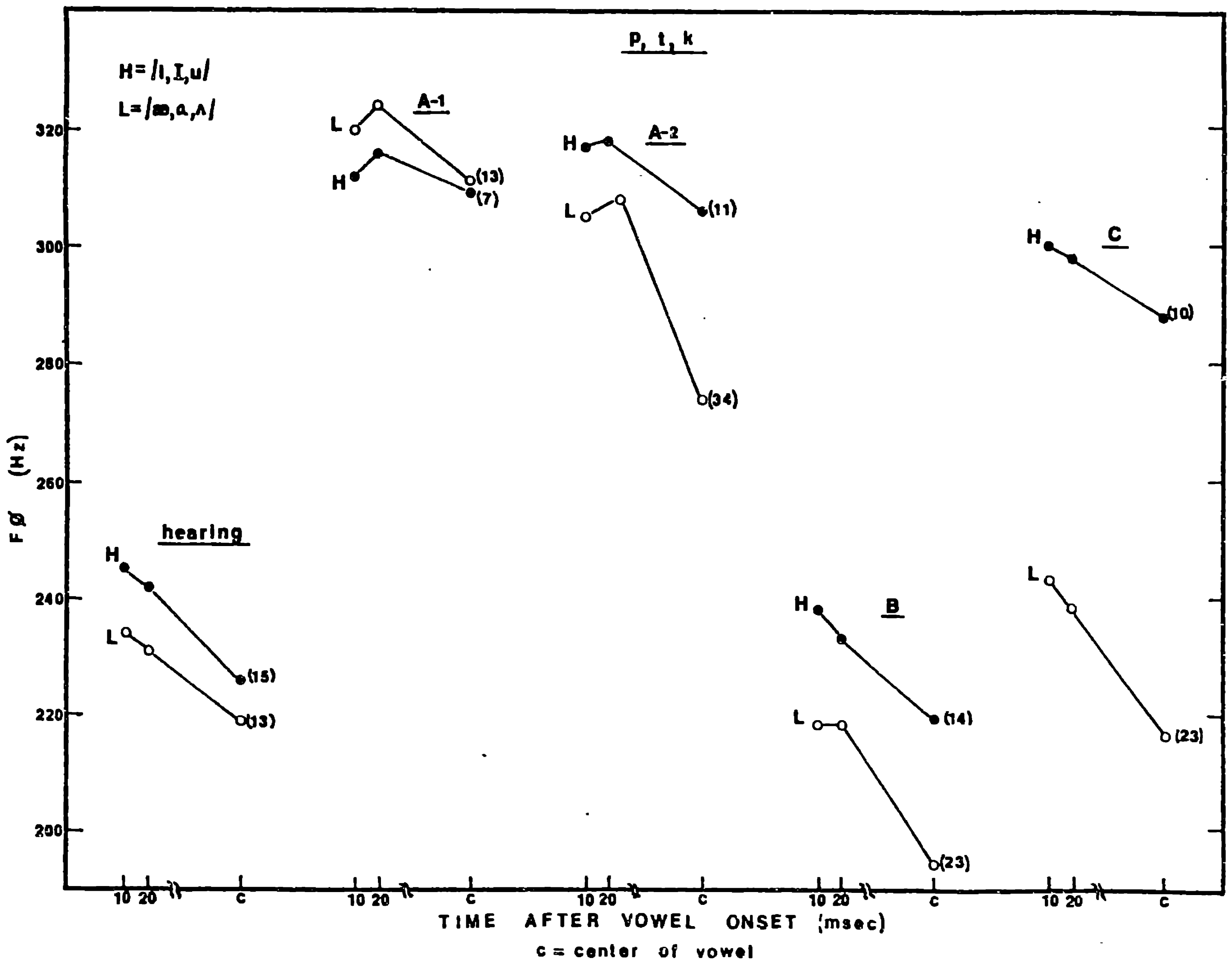


Figure 4.19 - Grouped F0 data for target nouns containing high (H) and low (L) vowels preceded by the voiceless stop consonants /p,t,k/. Numbers in parentheses represent averaged values of [F0(20) - F0(CV)]. (See text for explanation.)

for Group C).

The numbers shown in parentheses in Figure 4.19 represent the average difference between F0 near syllable onset and F0 at the center of the target vowel (i.e., the measure [F0(20) - F0(CV)]) for each group of speakers and for each set of target vowels. This measure is higher for nouns containing low vowels than for nouns containing high vowels for all groups of speakers except the hearing controls. (The difference in this measure for high versus low vowels is particularly large (i.e., 11 Hz versus 34 Hz) for the speakers of Group A-2.) Given that the low vowels in the utterances studied are preceded by the alveolar and velar consonants /t/ and /k/ (in the target nouns "top", "cat" and "cut"), these data would again seem to suggest that constraints on the position of the tongue near syllable onset may influence (i.e., increase) F0 for many of the deaf speakers.

In order to examine this possibility more carefully, F0 data were averaged across vowels for each place of consonant articulation for the four groups of deaf speakers and for the hearing controls. These averaged data, together with the corresponding values of [F0(20) - F0(CV)], are shown in Figure 4.20. (Numerical F0 data and the associated numbers of tokens are provided in Table A2.14 in Appendix 2.) As noted earlier, place of consonant articulation appears to have little influence on the amount by which F0 changes over the course of

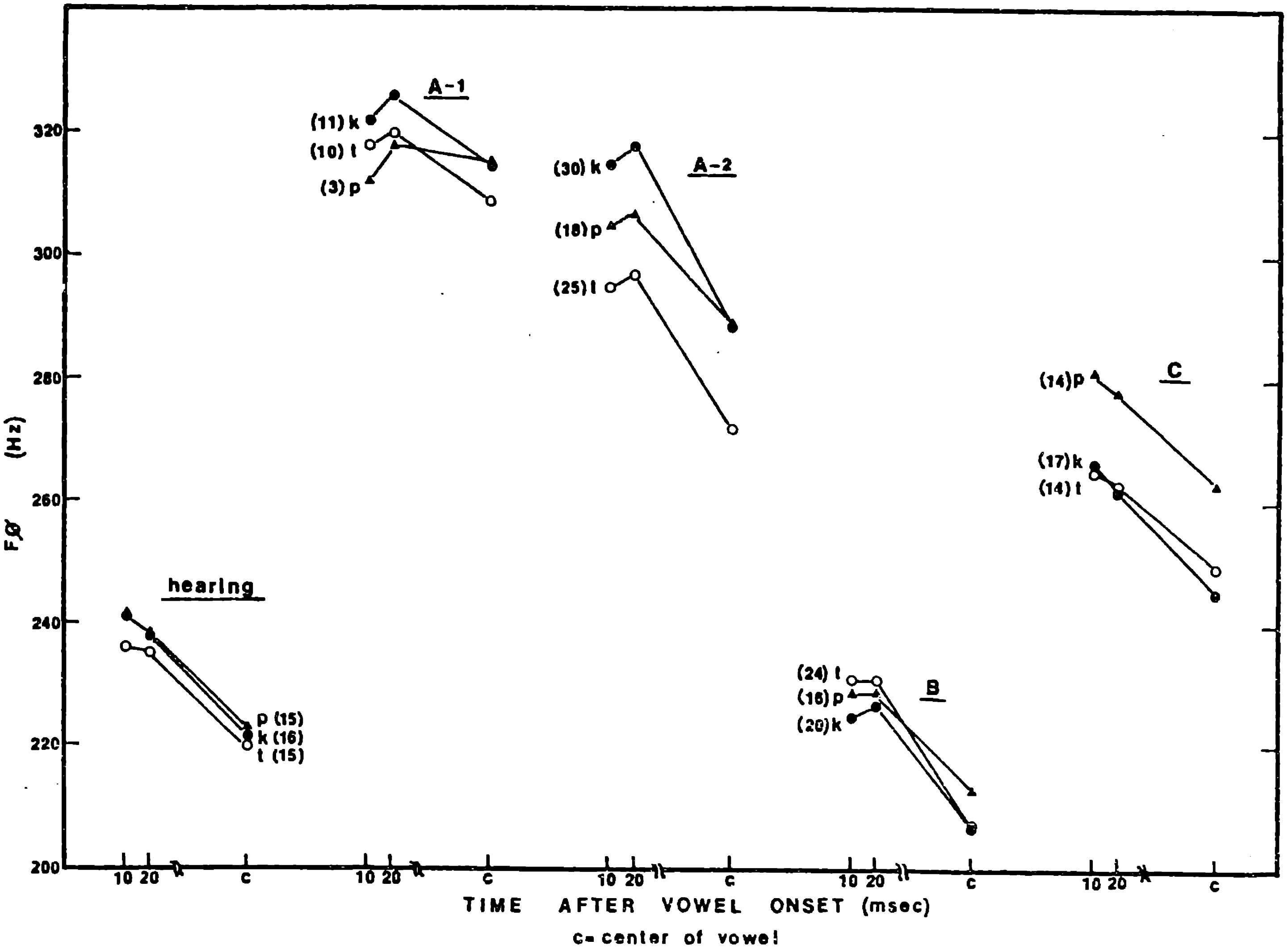


Figure 4.20 - Grouped F0 data for target nouns beginning with the voiceless stop consonants /p,t,k/. Numbers in parentheses represent averaged values of [F0(20) - F0(CV)]. (See text for explanation.)

the target syllables for the hearing speakers, and the same appears to be true for the deaf speakers of Group C. (The implications of this latter observation for an explanation of the excessive vowel-to-vowel variations in F0 produced by the speakers of Group C will be considered in Section 4.2.2.)

For the remaining three groups of deaf speakers, (particularly for Groups A-1 and A-2), F0 changes more over the course of syllables beginning with the alveolar and velar consonants /t/ and /k/ than with the labial consonant /p/. (Note, for example, for the speakers of Groups A-1 and A-2, that, although F0 is comparable near the center of vowels beginning with the consonants /p/ and /k/, F0 is somewhat higher near syllable onset for the latter consonantal context.) As will be discussed more fully in Section 4.2.2, these data again suggest that tongue height is relatively more important than jaw position in determining the vowel-related variations in F0 produced by the majority of deaf speakers (assumptions 2 and 3, Section 4.1.1).

4.2 Discussion of Mechanisms

4.2.1 Acoustic Coupling

In Section 1.3, two hypotheses were made about how acoustic coupling between the vocal folds and the supralaryngeal vocal tract might contribute to the exaggerated vowel-related variations in F0 produced by some deaf speakers. The first of these hypotheses was that an inappropriate laryngeal posture or mode of vocal-fold vibration (e.g., that associated with breathy voice quality) might be particularly sensitive to impedance changes in the supraglottal system, resulting in an increased coupling effect of a low-frequency first formant on F0. The second hypothesis was that acoustic coupling between F0 and F1 might lead to large upward breaks in F0 (or to instability in voicing) in cases for which the frequencies of the first formant and the fundamental were comparable and vocal-tract damping was relatively low. (See Section 1.3 for a discussion of factors motivating these two hypotheses.)

The data collected in the present study provide little support for an acoustic-coupling mechanism for (exaggerated) vowel-related F0 change. While the deaf speakers often produced vowels with low-frequency first formants (e.g., /i,u/) with inappropriately high F0, an inverse correlation between F0 and F1 was not observed for voiced consonantal segments. In particular, F0 on both nasal consonants and sonorants appeared to depend less on the presence of a low-frequency first formant than on tongue height for the adjacent target vowel (Sections 4.1.2 and 4.1.3). In most instances, F0 on the nasal or sonorant segment was lower than F0 on the vowel, even when the vowel was characterized by a relatively high-frequency F1.

A close interdependence between F0 on syllable-initial nasal consonants and the following stressed vowel has been noted for hearing adults by Ewan [Ewa79a]. In his study, male and female speakers repeated the target utterances /umu/ and /umɑ/ (in the context "Say ___ again."), and F0 measurements were made at the center of the nasal consonant and at a point approximately one hundred milliseconds into the following vowel. Statistical analyses of these data showed that, for a given vowel context, the difference between F0 on the consonant and F0 on the vowel was not significant; F0 was, however, significantly higher on /m/'s preceding the high vowel /u/ than on /m/'s preceding the low vowel /ɑ/. Ewan concluded that F0 on the consonantal segments had been determined by coarticulatory anticipation of the tongue and jaw for the following vowel (rather than by the presence of a low-frequency F1) and used these results to argue against an acoustic-coupling mechanism for (vowel-related) F0 change.

On syllable-initial nasals (and sonorants), F0 was relatively independent of vowel context for the hearing boys and girls in the present study, and F0 on these consonants tended to be higher than F0 on the following target vowel. While these observations would appear to be consistent with an acoustic-coupling hypothesis, it is perhaps more likely that they reflect the phrase-final position of the target syllables in the test utterances (i.e, a phrase-final fall in F0). The observation that F0 on nasals in syllable-final position tended to be lower than F0 on the target vowel for the hearing speakers (Figure 4.3) would support this latter explanation.

Examination of the F0 data for syllables containing nasals and sonorants showed that the deaf boys and girls sometimes produced abrupt upward changes in F0 or breaks in voicing in the vicinity of consonant-vowel or vowel-consonant boundaries (Sections 4.1.2 and 4.1.3). However, such changes in the F0 contours were not consistently related to downward changes in the frequency of F1 (i.e., to the introduction of a supraglottal resonance at a frequency near F0), as would be predicted by the second of the two acoustic-coupling hypotheses outlined above. As noted earlier, some of these changes in the F0 contours may have been associated with the insertion of stop consonants between the vowel and nasal (or sonorant) segments or with some other inappropriate concatenation of adjacent speech sounds.

Finally, a positive correlation between the amount of vowel-related variability in F0 used by the deaf speakers and spectral measures indicative of incomplete laryngeal closure during voicing (i.e., an overly high-amplitude first harmonic or spectral noise at high frequencies) was not observed (Section 3.2). The lack of such a correlation (together with the lack of a consistent inverse relationship between F0 and F1) suggests that an inappropriate laryngeal posture, such as that associated with breathy voice quality, does not contribute appreciably to an increased coupling effect of F1 on F0 and, thus, to the large vowel-related variations in F0 produced by many of the deaf speakers (see Section 1.3 and above).

4.2.2 Vocal-Fold Tension

The segmental variations in F0 produced by the deaf boys and girls in the present study appear to be better explained by a mechanism (or mechanisms) which assumes that they result from changes in vocal-fold tension associated with articulatory maneuvers (and, perhaps, with articulatory postures) used in vowel and consonant production. In particular, the data described earlier in this chapter (Sections 4.1.2 through 4.1.4) suggest that changes in the position of the tongue body may have a major influence on the F0 of some deaf speakers.

For each of the consonantal contexts examined in the study, a close relationship was observed between F0 on the consonant (or F0 just after consonant release) and the height of the adjacent target vowel. F0 on (or just after) consonants adjacent to low vowels tended to be lower than on consonants adjacent to high vowels, independent of constraints on the frequency of the first formant (e.g., for nasals) or on jaw position (e.g. for labial stops). Furthermore, for the majority of deaf speakers, a greater amount of F0 change was imposed on syllables for which the position of the tongue body was constrained at syllable onset (e.g., for syllables beginning with the alveolar and velar consonants /t/ and /k/) than on syllables for which the tongue was relatively free to anticipate a configuration appropriate to that for the following vowel (e.g., for syllables beginning with the labial consonant /p/).

The remainder of this section will consider more carefully (through a discussion of anatomical and physiological data reported in the literature for hearing speakers) how such changes in tongue position might influence vocal-fold tension, and will examine factors which may exaggerate the resulting changes in F0 in the speech of some deaf individuals. In line with this latter objective, differences and similarities in the F0 and formant data for the various groups of speakers described in Chapters 3 and 4 will be considered and reference will be made to the "vocal-fold tension" hypotheses discussed in Section 1.2 and 1.3.

Figure 4.21 shows a schematic representation of the anatomy relevant to the present discussion, including the larynx, the hyoid bone and several muscles whose activity is known to be correlated with changes in F0 (see below). This figure is taken from a recent paper by Honda [Hon81], which attempts to explain the vowel-related variations in F0 produced by hearing speakers in terms of anterior-posterior and rotational movements of the hyoid-larynx complex. More specifically, Honda has argued that shifting the tongue root forward for the production of high vowels, such as /i/ and /u/, will also cause the hyoid bone to move forward and to tilt the thyroid cartilage anteriorly (at the lateral thyrohyoid ligament), resulting in an increased longitudinal tension on the vocal folds and, thus, an increase in F0.

Honda's argument was originally motivated by electromyographic and cinefluorographic studies which showed that increases in F0 were associated both with activity in the geniohyoid muscle [EriLib77] and with forward movements of the hyoid bone [Sap78]. The geniohyoid is an anterior suprahyoid muscle which extends from the interior surface of the mandible to the anterior body of the hyoid (Figure 4.21) and which

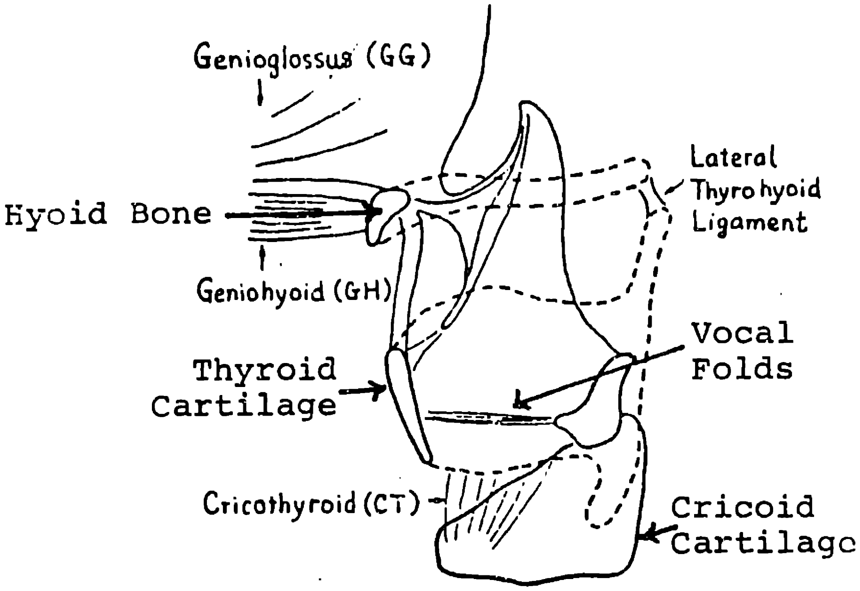


Figure 4.21 - Schematic illustration of the anatomy of the hyoid-larynx complex. (After Honda [Hon81]; used with permission.)

contracts to move the hyoid bone forward when the mandible is fixed [Har76]. Because activity in the geniohyoid was also correlated with activity in the *cricothyroid* muscle (which contracts to create a change in the angle between the thyroid and cricoid cartilages) [EriLib77], Honda reasoned that the two muscles acted together to tilt the thryoid forward and, thus, to increase F0. He also reasoned that a similar phenomenon (i.e., a forward shift of the hyoid and an anterior tilting of the thyroid) might be associated with contraction of the *posterior genioglossus* muscle during the production of high vowels. The posterior genioglossus, which forms much of the central core of the tongue and attaches directly to the anterior surface of the hyoid bone (Figure 4.21), contracts to move the root of the tongue forward and thus to force the tongue body upwards.

In his own study, Honda [Hon81] examined measures of hyoid movement (using an optical tracking system) and of EMG activity in the geniohyoid, cricothyroid and posterior genioglossus muscles as a function of changes in F0 and in vowel quality. These data were collected during the production of Japanese two-mora nonsense words (e.g., /ɑɑ/, /ii/, /iɑ/, /ɑmi/) with rising, falling and steady pitch-accent patterns. Measures of posterior genioglossus activity were also compared with F0 data (collected by Lehiste and Peterson [LehPet61]) for a set of American English vowels.

The results of Honda's analyses were consistent with those of the earlier studies [Sap78];[EriLib77], in that high F0 (in rising or falling pitch accents) was correlated both with increased activity in the geniohyoid and cricothyroid muscles and with a more forward position of the hyoid bone. Furthermore, the overall pattern of activity in the geniohyoid muscle was comparable to that in the cricothyroid, providing some support for Honda's argument that forward movements of the hyoid (such as those associated with contraction of the geniohyoid) aid in tilting the thyroid cartilage forward during the production of high F0.

The results of Honda's analyses were also consistent with his speculations concerning the relationship of tongue movement to hyoid movement and to vowel-related F0 change in that: 1) a more forward position of the hyoid bone, as well as greater activity in the posterior genioglossus muscle, was associated with the production of the high vowel /i/ than with the low vowel /ɑ/ (in two-mora nonsense words with level pitch accents); and 2) changes in the level of posterior genioglossus activity during the production of the American English vowels were closely correlated with vowel-to-vowel variations in F0. The latter phenomenon is illustrated in Figure 4.22, which is also taken from Honda's paper.

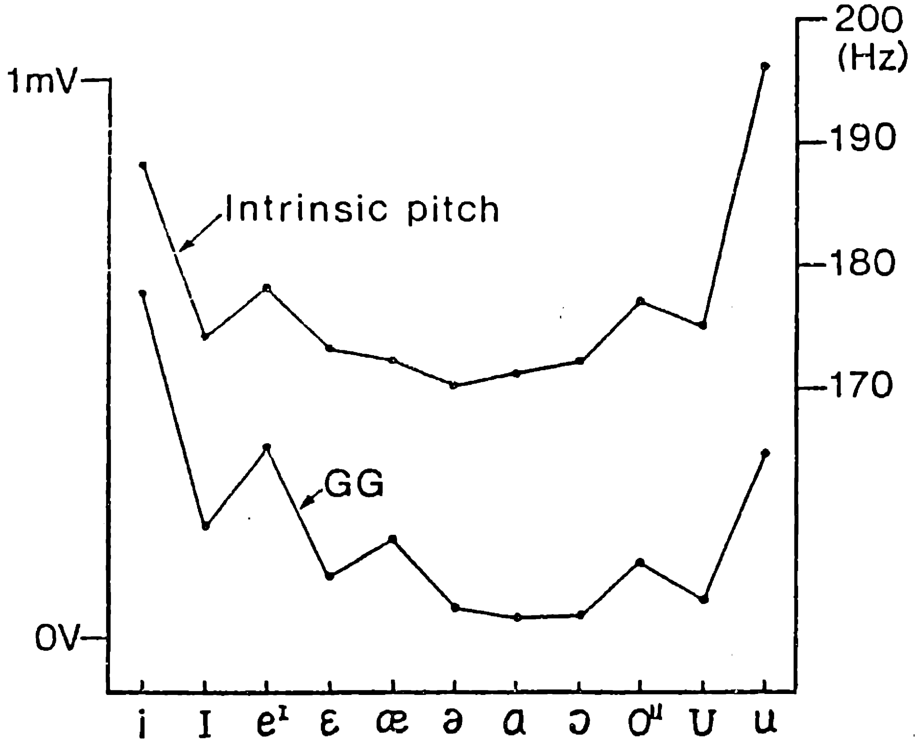


Figure 4.22 - Vowel-to-vowel variations in F0 and in EMG activity of the posterior genioglossus muscle for hearing adult speakers. (After Honda [Hon81].) The F0 data are taken from Lehiste and Peterson [LehPet61] for vowels preceded by the stop consonants /p,t,k/. The EMG data, collected in an experiment at Haskins Laboratories, represent peak activity of the posterior fibers of the genioglossus for utterances of the form /əpVp/, as produced by a native speaker of American English [Hon81]. (Used with permission.)

While Honda argues that the major influence of tongue and hyoid movement on F0 is accomplished through a forward rotation of the thyroid cartilage and an associated increase of longitudinal tension on the vocal folds, he notes that changes in the position of these structures (i.e., the tongue and hyoid) might also affect vertical tension in the larynx and, thus, F0. As discussed in Section 1.3, Ohala has suggested that interactions between tongue height and vertical tension in the larynx may play an important role in determining vowel-related variations in F0 [Oha77];[Oha73], but the exact nature of this (hypothetical) phenomenon is still not well understood.

The question of interest in the present discussion is, of course, whether the mechanism proposed by Honda will prove useful in explaining the F0 data described in Chapters 3 and 4, particularly: 1) the exaggerated vowel-related changes in F0 produced by many of the deaf boys and girls; and 2) the relatively high F0 following the release of alveolar and velar stop consonants (relative to F0 at the center of the following vowel) observed for the majority of deaf speakers.

With respect to the latter phenomenon, several observations suggest that this question may be answered in the affirmative. First, cinefluorographic data reported by Perkell [Per69] show that the width of the pharynx (more specifically, the distance between the dorsum of the tongue and the cervical vertebrae C2 and C3) is relatively large during the closure period and just after the release of the stop consonants /t/ and /k/, suggesting (at least for the vowel contexts examined) that the position of the tongue body is fairly far forward (and high) for both alveolar and velar articulations. Since the posterior fibers of the genioglossus would play a major role in positioning the tongue body in this manner [Har76]; [Per69], it seems likely that the position of the hyoid (and possibly F0) would be affected. Furthermore, forward movement of the tongue body during the production of such consonants may be aided by the contraction of other suprahyoid muscles, such as the mylohyoid, the stylohyoid and the anterior belly of the digastric, which also act to draw and/or to tilt the hyoid bone, and, thus, the thyroid cartilage, forward [Har76].

Perkell's data also indicate that the width of the pharynx during the articulation of the labial stop consonant /p/ (in utterances of the form /hə'Cɛ /) is relatively more dependent on pharynx width during the following stressed vowel than it is during the articulation of the alveolar and velar stops /t/ and /k/. Perkell attributes this result to the lack of a primary role for the tongue during the production of the /p/ and, thus, to coarticulatory anticipation of the tongue for the following vowel. If F0 were influenced by movements of the tongue body (through horizontal changes in hyoid position) as Honda has argued, one might expect changes in F0 over the course of consonant-vowel syllables to be somewhat less dependent on vowel height for syllables beginning with labial stop consonants (by virtue of such coarticulatory phenomena) than for syllables with alveolars or velars. One might also expect such changes in F0 to be greatest in magnitude for syllables in which a low back vowel such as /ɑ / was preceded by an alveolar or velar stop. Such expectations are consistent with the F0 data obtained for the majority of deaf speakers in the present study (Sections 4.1.4 and 4.1.5).

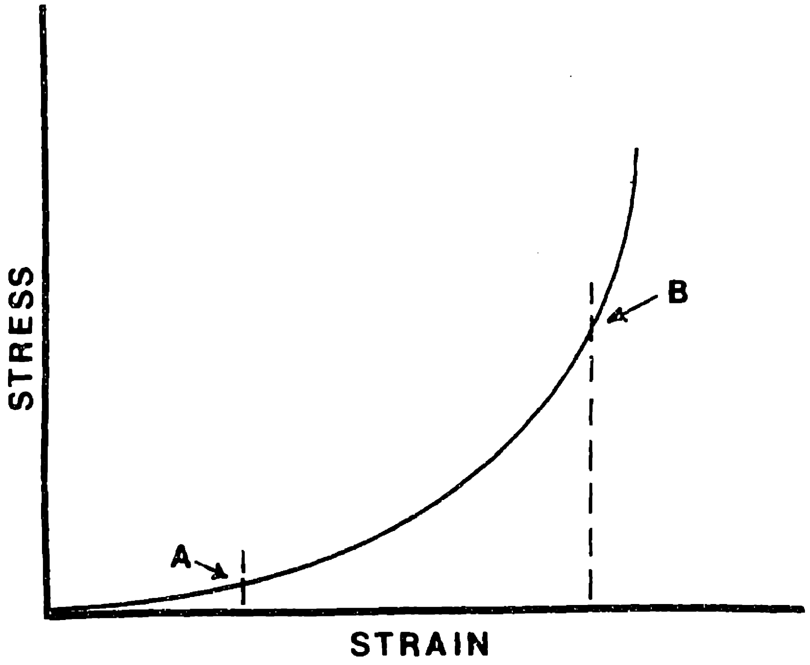
One issue which needs to be addressed, however, is why place of stop consonant articulation apparently does not influence the amount of F0 change imposed on the F0 contours produced by the hearing speakers (or by the deaf girls in Group C). While a definitive answer to this question is impossible on the basis of the data collected in this study, a number of speculations can be made. For example, one possibility is that the articulatory maneuvers used by most (but not all) of the deaf speakers in producing the stop consonants /t/ and /k/ are somewhat more extreme than those used by the hearing boys and girls [Mar68], resulting a greater degree of tension in those muscles which influence hyoid (and larynx) position and thus, presumably, F0.

A second possibility is that an articulatory configuration for the target vowel is achieved more quickly by the hearing speakers (e.g., during the aspiration period for the voiceless stops /t/ and /k/) and, thus, that tongue position (and F0) after vowel onset is relatively less dependent upon consonantal context for the hearing than for the deaf boys and girls. Numerous investigators have noted that the production of both vowel and consonantal segments is often prolonged in the speech of the deaf (e.g.,[Lin76];[Nic75]), and it is conceivable that the same may be true of transitions between the two types of speech sounds.

With respect to the first of the two questions posed above, it is possible that the mechanism for vowel-related F0 change proposed by Honda [Hon81], together with the results of a study of length-tension characteristics of vocal folds by Kakita, Hirano and Ohmaru [KakHir81], explain at least some of the exaggerated vowel-to-vowel variations in F0 produced by the deaf speakers (i.e., those produced by the deaf boys and girls in Group A-2). In their study, Kakita, et al. excised larynges from normal adult dogs and measured the stress-strain characteristics of the vocal fold tissue. This was done by fixing the prepared larynges at one end (i.e., at the thyroid cartilage) and hanging them vertically. Known weights were then attached to the other end (i.e., to the arytenoid cartilage) and the resulting deformations in vocal-fold length were measured.

A schematic representation of the results of these experiments is shown in Figure 4.23, as are the equations used in calculating vocal-fold stress and strain. (Note that stress is proportional to vocal-fold tension while strain is proportional to length deformation.) For the purposes of the present study, the most important observation to be made from the curve shown in the top section of Figure 4.23 is that the stress-strain relationship for the vocal-fold tissue is non-linear. Thus, a given change in the length of the vocal folds will produce a greater change in vocal-fold tension (and, presumably, in F0) when the strain on the vocal folds (and, thus, F0) is already relatively high (point B in Figure 4.23) than when the strain on the vocal folds is low (point A).

Assuming that Honda's arguments are correct and that the vocal folds are stretched longitudinally during the articulation of high vowels, this non-linear stress-strain relationship would predict that larger increases in F0 should characterize vowels such as /i/ and /u/ when they are produced at high (as compared with low) F0 levels. Data on vowel-to-vowel variations in F0 as a function of tone (in Taiwanese Chinese) [Zee78] and of sentence position (in American English) [ShaPie79] appear to be consistent with this



$$\text{Stress} = \frac{T}{S_0}$$

$$\text{Stress} = \frac{\Delta L}{L_0}$$

where:

T = tension

S_0 = cross-sectional area (normal to the direction of the tension)

ΔL = elongation

L_0 = length at no load

Figure 4.23 - Stress-strain relationship for vocal-fold tissue. The curve in the top section of the figure is a schematic representation of data collected by Kakita, Hirano and Ohmaru [KakHir81]. See text for a description of points A and B.

prediction. In both situations, frequency differences between high and low vowels tend to be greater in magnitude when the overall F0 level is relatively high (e.g., for high tones and sentence-initial positions) than when F0 level is relatively low (e.g., low tones and sentence-final positions).

On the basis of the studies of Honda and of Kakita, et al., one would expect both mean F0 and articulatory skill to play a role in determining the amount of vowel-related F0 variability used by deaf (and hearing) speakers. More specifically, exaggerated vowel-to-vowel variations in F0 should be most common among deaf speakers who: 1) maintain an F0 level which is somewhat higher than normal for all vowels; and 2) produce extreme tongue displacements, and thus F1 and F2 values, appropriate to the articulation of high vowels such as /i/ and /u/. Such expectations appear to be compatible with the F0 and formant data obtained in the present study for the deaf speakers of Group A-2 (Section 3.1.3). (The relatively small vowel-related variations in F0 observed for deaf speakers whose mean F0 was close to normal (e.g, Group B) and for deaf speakers whose articulatory abilities, particularly with respect to the production of high vowels, were poor (e.g., Group A-1) are also compatible with these expectations.)

The vowel diagrams presented in Section 3.1.3 indicate that the deaf boys and girls in Group A-2 also made use of a range of F1 values which was considerably wider than that used by the hearing controls (Figure 3.6). In particular, these speakers tended to produce the low vowels /æ , ɑ , ʌ / with higher-than-normal F1 and, thus, presumably, with greater-than-normal jaw opening. While jaw position per se did not appear to have a dominant influence on the F0 contours produced by the speakers in Group A-2 (see the data for syllables with labial consonants presented in Secions 4.1.4 and 4.1.5), this exaggerated amount of F1 variability could reflect a more general use of extreme articulatory habits by these boys and girls. Such extreme articulation might lead to a greater range of values of tension in those muscles which influence hyoid/larynx position and, thus, to increased (vowel-related) variations in F0.

The exaggerated vowel-to-vowel variations in F0 produced by the deaf girls in Group C are somewhat more difficult to explain in terms of changes in articulatory configuration (i.e., changes in tongue height) and/or mean F0. In contrast to the other three groups of deaf speakers, the F0 contours produced by these girls were relatively uninfluenced by the constraints on tongue position associated with the production of alveolar and velar stop consonants (Section 4.1.5 and Figure 4.20), and differences in the F0 levels and amounts of formant variability used by the four girls were large (Section 3.1.3). Despite these inconsistencies, however, each of the girls produced the vowel /ɑ/ with an F0 comparable to normal and the vowels /i/ and /u/ with excessively high F0.

As discussed in Section 3.1.3, it is possible that age-related factors (i.e., adolescent voice change or similarities in speech training) may have played a role in determining the large vowel-related variations in F0 produced by this group of speakers. Another possibility (also potentially related to speech training) is that these girls have learned to associate an overall increase in articulatory tension, but not necessarily a high and fronted position of the tongue body, with the production of high vowels such as /i/and /u/, and a more relaxed articulatory posture, together with a low tongue and jaw position, with the production of low vowels such as /ɑ/. (Note that, while only one of the girls in group C produced the low F1's appropriate to the vowels /i/ and /u/, all four girls produced high F1's for the vowel /ɑ/ (Table A1.2).)

Before the present discussion is concluded, two other aspects of the F0 and formant data obtained for the deaf speakers in the present study should be noted and their relevance to hypotheses discussed in Section 1.3 should be considered. The first of these aspects is the apparent relationship among an overly high mean F0, low F0 variability and limited formant variability (particularly the limited range of F2 values) observed for the speakers of Group A-1 (Section 3.1.3).

As noted in Section 1.3, the production of a narrow range of F2 values may be one consequence of an articulatory posture in which the tongue is held in a low and backed position relative to that used by hearing speakers [SteNic79], resulting in a greater-than-normal degree of pharyngeal constriction during vowel production. Ling [Lin76] has suggested that such excessive pharyngeal constriction (or "pharyngeal tension") may serve to induce an inappropriate degree of tension in the larynx itself, thus limiting F0 variability (and, presumably, increasing mean F0) in the speech of some deaf individuals [Lin76].

While the data for the speakers of Group A-1 appear to be compatible with Ling's hypothesis, the way in which high laryngeal tension might act to limit (vowel-to-vowel) variability in F0 is not clear, particularly in light of the mechanism for exaggerated vowel-related F0 change proposed earlier in this section. An alternative possibility (which is more consistent with the arguments presented above) is that, <u>in addition to</u> increasing tension in the larynx, pharyngeal constriction serves to limit the amount of front-back (and perhaps up-down) tongue movement a deaf speaker can achieve [SteNic79], thus resulting both in poor vowel articulation and in reduced vowel-to-vowel variations in F0. The articulation of consonantal segments, such as the alveolar /t/, might still be possible, however, since movements of the tongue tip (or the front part of the tongue body) would be relatively less constrained by pharyngeal constriction than would movements of the tongue root.

The second aspect of the data which deserves mention is the amount of vowel-related F0 variability used by the deaf girls of Group B. As noted in Section 3.1.3, these girls were virtually indistinguishable from the hearing controls with respect to both mean F0 and articulatory skill (as evidenced by the range of F1 and F2 values produced), but still produced vowel-to-vowel variations in F0 which were somewhat larger than those observed for the hearing speakers. One possibility suggested by these data (and discussed more fully in Section 1.3) is that relatively large changes in F0 are, in fact, a natural consequence of vowel articulation, for which hearing speakers learn to compensate (e.g., through adjustments in laryngeal tension or position).

The results of a study by Eguchi and Hirsh provide some indirect support for this latter hypothesis. These investigators measured F0 for the six vowels /i,ɛ,æ,ɑ,ɔ,u/ in sentence context (five repetitions per speaker) and examined intraspeaker standard deviations in F0 as a function of age [EguHir69]. (F0 measurements were made for groups of speakers ranging in age from three to thirteen years.) They found that, on average, intraspeaker variability in F0 decreased with age until a minimum was reached at approximately ten to twelve years, a result which held both for the absolute magnitudes of the standard deviations and for the standard deviations normalized by the corresponding mean F0's.

Eguchi and Hirsh did not present F0 data for individual vowel contexts, and thus it is impossible to know the extent to which the standard deviations which they reported resulted from vowel-to-vowel (as opposed to cross-repetition) variations in F0. Nonetheless, as noted by Kent [Ken76], the results of this study suggest that accuracy of laryngeal control during vowel articulation may improve continuously for hearing speakers over a period of at least seven to nine years. If Kent's interpretation is correct, and if this improvement is dependent to a considerable extent on auditory feedback, it is conceivable that the resulting (vowel-related) stability in F0 observed for hearing speakers may be delayed for (or never achieved by) many deaf individuals. This possibility, however, like many of the others discussed in this section, remains to be tested with additional data.

CHAPTER 5

CONCLUDING COMMENTS

This chapter summarizes the major results of the thesis and considers the implications of these results for speech training and for future research efforts. Some limitations of the analyses performed in the study are also considered.

5.1 Summary of Results

5.1.1 Vowel Articulation and Fundamental-Frequency Control

The first major objective of this thesis was to document more carefully the relationship between the vowel-to-vowel variations in F0 produced by deaf speakers and their proficiency at vowel articulation. This goal was pursued through a comparison of F0 and formant-frequency data for groups of deaf and hearing boys and girls and through an examination of these data relative to intelligibility scores for the deaf speakers.

For the majority of deaf boys and girls in the study, a close relationship was observed between vowel-related variability in F0 and articulatory skill. In general, greater F0 variability was observed for deaf speakers who produced a relatively wide range of vowel sounds (i.e., of F1 and F2 values) than for speakers whose articulatory capabilities were more limited. Intelligibility scores also tended to be higher among those deaf speakers who produced larger vowel-to-vowel variations in F0.

Although the amount of F0 variability used by the deaf boys and girls was often much larger than that used by the hearing controls, the direction in which F0 changed as a function of vowel height was comparable for both deaf and hearing speakers. In both cases, F0 was higher for high vowels, such as /i/ and /u/, than for low vowels, such as /ɑ/.

Based on these observations, it was concluded that the vowel-to-vowel variations in F0 produced by the deaf speakers were qualitatively, if not quantitatively, like those produced by the hearing controls (i.e., in that they were, in some way, a consequence of articulatory maneuvers used in vowel production). The data were also taken to refute the hypothesis, set forth by a number of previous investigators [Hor77];[AngKop64], that deaf speakers use excessive vowel-related F0 variability rather than articulatory (or formant) variability as a means of distinguishing one vowel sound from another.

5.1.2 Mechanisms for Exaggerated F0 Change

A second and more general objective of this thesis was to explore mechanisms which might account for the exaggerated vowel-related variations in F0 observed for many deaf speakers. In line with this objective, two mechanisms proposed to account for normal interactions between vowel articulation and F0 were reviewed: 1) *acoustic coupling* between the vocal folds and the lowest resonance (or formant) of the supralaryngeal vocal tract; and 2) changes in *vocal-fold tension* associated with tongue, jaw and/or larynx maneuvers used in vowel (and consonant) production.

A number of hypotheses were then made about how certain aspects of the speech of the deaf (e.g., inappropriate laryngeal posture or extreme articulatory habits) might influence the operation of these two mechanisms, resulting in larger-than-normal vowel-related variations in F0. Finally, F0, formant and spectral data for relevant vowel and/or consonantal contexts were examined (for the same boys and girls as above), and the results of these analyses were interpreted relative to the hypothetical mechanisms for exaggerated F0 change.

The data collected in the study provided little support for an acoustic-coupling mechanism. While vowels with low-frequency first formants were often produced with inappropiately high F0 by the deaf speakers, F0 on nasal consonants and sonorants appeared to depend less on the presence of a low-frequency F1 than on tongue height for the adjacent target vowel. The lack of a positive correlation between spectral measures indicative of incomplete laryngeal closure during voicing and excessive vowel-related variability in F0 also argued against one of the acoustic-coupling hypotheses (i.e., the hypothesis that a "spread" laryngeal posture, such as that associated with the production of breathy voice quality, might lead to an increased coupling effect of F1 on F0).

The segmental variations in F0 produced by the majority of deaf speakers appeared to be better explained by a vocal-fold tension mechanism--in particular, an extension of a mechanism proposed by Honda [Hon81] to account for normal interactions between vowel articulation and F0. To summarize, Honda's mechanism assumes that shifting the tongue root forward for the production of high vowels will also cause the hyoid bone to move forward and to tilt the thyroid cartilage anteriorly. This rotation of the thyroid cartilage results in an increased longitudinal tension on the vocal folds and, thus, an increase in F0. On the basis of the non-linear nature of the stress-strain relationship for vocal-fold tissue, it was argued that such increases in vocal-fold tension may be somewhat greater in magnitude when the tension on the vocal folds (and, thus, mean F0) is already relatively high, resulting in somewhat larger increases in F0 during the articulation of high vowels. This argument appears to be consistent with F0 data reported in the literature for hearing speakers as well as with the following observations made in the present study:

1) Exaggerated vowel-to-vowel variations in F0 were produced by deaf speakers who maintained a mean F0 which was somewhat higher than normal, and who were capable of articulatory configurations appropriate to high vowels. Furthermore, the amount of F0 variability used by these speakers was determined primarily by an excessively high F0 for the high vowels /i,I,u/ relative to F0 for /ɑ/.

2) Smaller vowel-to-vowel variations in F0 were produced by deaf speakers whose mean F0 was comparable to normal and by speakers whose articulatory skills, particularly with respect to the production of high vowels, were poor.

The exaggerated vowel-related variations in F0 produced by a smaller group of deaf speakers (four girls) were more difficult to explain on the basis of the mechanism described above. While all four speakers produced the vowel /ɑ/ with an F0 comparable to normal and the vowels /i/ and /u/ with excessively high F0, they differed considerably among themselves with respect to mean F0 and articulatory skill. Each of the girls was, however, approximately fourteen years old, suggesting that age-related factors (e.g., adolescent voice change or similarities of speech training) may have contributed to their problems with vowel-related F0 control.

5.2 Implications for Speech Training

The results of this study have a number of implications for speech training. For example, the inverse relationship between (vowel-related) F0 variability and formant variability observed for the majority of deaf boys and girls suggests that one means a deaf speaker can use to eliminate inappropriately large variations in F0 from his or her speech is to produce a smaller range of articulatory maneuvers. Unless the teacher of the deaf is aware of this potential interaction between articulation and F0, any attempt to train more stable fundamental-frequency control could inadvertently lead to reduced speech intelligibility. Furthermore, the design and use of visual or tactile aids which provide only a single channel of information (e.g., a display of an F0 contour) could prove inappropriate, unless an attempt is made to supplement this information during speech training (e.g., by simultaneously attending to the adequacy of segmental articulation).

The relevance of both of these cautions is suggested by the results of a study performed by Nickerson, Stevens and Rollins [NicSte79b]. In this study, intelligibility data were collected for a group of deaf speakers before and after training with a visible F0 display, and were examined relative to objective measures of the speakers' proficiencies at F0 control. After training, several of the deaf speakers showed substantial reductions in the number of vowels which they produced which deviated by more than 50 Hz from an idealized (sentential) F0 contour. At the same time, however, a slight <u>decrease</u> was observed in the average intelligibility of unfiltered sentence material for these speakers. Nickerson, et.al. reasoned that the lowered intelligibility scores might have been the result of modifications in vowel articulation which the deaf speakers made in order to avoid jumps in pitch associated with certain articulatory configurations. (As noted in Section 1.1, Nickerson, et.al. also found a positive correlation between intelligibility scores for a larger group of deaf speakers and the amount of deviation from their idealized F0 contour, again suggesting a relationship between good articulation and excessive variability in F0.)

The large vowel-related variations in F0 produced by at least some of the deaf speakers in this study appeared to be related not only to articulatory (i.e., formant) variability but also to the maintainance of a mean F0 which was somewhat higher than that used by the hearing controls. On the other hand, the use of a mean F0 which was <u>extremely</u> high relative to normal was associated both with limited F0 variability and with the production of a greatly reduced range of F1 and F2 values. Taken together, these two observations suggest that a first priority in simultaneously training good articulation <u>and</u> stable fundamental-frequency control may be to work toward establishing a mean F0 appropriate to a deaf speaker's age and sex.

A second priority, given that many of the deaf speakers who produced exaggerated variations in F0 also produced a wider-than-normal range of F1 values, might be to discourage the use of an excessive amount of jaw movement during the production of low vowels such as /ɑ/ or /æ/. While jaw position per se did not appear to have a dominant influence on the F0 contours produced by the deaf speakers, this exaggerated amount of F1 variability could reflect a more general use of extreme articulatory habits by the deaf boys and girls. As discussed in Section 4.2.2, such extreme articulation might lead to a greater degree of tension in those muscles which influence laryngeal tension and, thus, contribute to the prouction of increased vowel-related variations in F0.

Traditionally, the view has been held that lowering mean F0 to an appropriate level is desirable in speech training because it makes the deaf speaker's voice more pleasant to listen to and/or because it allows for a greater degree of F0 movement for use in intonation [BBN77];[Lin76];[Cal61]. The results of the present study, however, suggest that high F0 also tends to be associated with the production of a limited range of F1 and F2 values. To the extent that both of these problems reflect the use of (the same) incorrect speech posture (e.g., an overly constricted pharynx), correcting the use of an excessively high mean F0 may also improve the deaf speaker's ability to produce an appropriate range of vowel sounds. This observation provides one example of the way in which the development of good suprasegmental or postural skills can aid in the production of <u>segmental</u> features crucial to speech intelligibility [SteNic79].

Finally, while it may be convenient or necessary to distinguish among various aspects of speech production in research and in training that is concerned with the speech of the deaf, it is also important to recognize that these aspects are often not controlled independently of one another. The close interactions between articulation and F0 observed in this study stress the integral nature of the speech-production process. They also suggest the dangers of concentrating <u>solely</u> on one problem (or skill) during speech-training sessions, without at the same time monitoring the effects of this skill on other aspects of speech production.

5.3 Limitations of the Study and Directions for Future Research

The analyses in the present study were limited primarily by two factors: 1) the relatively small amount of data available for the various subgroups of deaf speakers; and 2) the need to make inferences about speech mechanisms solely on the basis of acoustic measurements. Both of these factors were, to some extent, a consequence of the exploratory nature of the thesis. Previous studies describing interactions between segmental articulation and F0 control by deaf speakers made use almost exclusively of averaged data, making it impossible to predict the types of and the extent of interspeaker differences which might be observed. Similarly, no good objective data were available to justify or to direct the use of more sophisticated techniques for examining laryngeal and articulatory control. The results of this study help to remedy both needs and, in so doing, suggest a number of areas for future research.

At the most general level, it would be useful to repeat many of the acoustic analyses described in the thesis with larger and different populations of deaf speakers and with a more extensive corpus. Examining data for students in other schools (both oral and non-oral), for example, would help to determine the extent to which the phenomena observed in this study were the result of peculiarities of speech training. Data for connected utterances and for spontaneous speech would provide useful information about interactions between articulation and F0, and their influence on speech intelligibility, in situations in which more meaningful communication is necessary.

Several specific lines of research are also suggested by the results of this study. One line would involve the collection of F0 data for different vowels, as a function of age, for deaf and hearing speakers. It was suggested earlier in the thesis, for example, that hearing speakers may learn to compensate (through changes in laryngeal tension and position) for relatively large changes in F0 which are a natural consequence of vowel articulation (Section 4.2.2). A comparison of developmental changes in the amount of vowel-to-vowel variability in F0 used by young (deaf and hearing) children would be useful in testing this hypothesis. Similarly, an examination of changes in F0 variability in the years surrounding puberty would be useful in gaining some insight into the effects of rapid laryngeal growth on vowel-related fundamental-frequency control.

A second line of research would be to examine more carefully interactions between mean F0 and vowel-related variability in F0--as a first attempt at testing the mechanism for exaggerated F0 change proposed above. Hearing speakers, for example, might be asked to phonate vowels at different pitch levels (without the benefit of auditory feedback), in order to determine whether increased vowel-to-vowel variations in F0 are consistently related to the use of an overly high mean F0. A more careful survey of intrinsic pitch phenomena in tone languages might also provide information relevant to this issue.

A third line of research would be to apply more sophisticated methods of analyzing laryngeal and articulatory control to smaller (select) populations of deaf speakers. For example, optical tracking methods, such as those employed by Honda [Hon81], might be used to determine the extent of front-back (or up-down) hyoid movement associated with the production of exaggerated vowel-to-vowel variations in F0. Correlation of such measurements with data obtained using more intrusive procedures, such as fiberoptic examination of vocal-fold length or electromyography of relevant intrinsic and extrinsic laryngeal muscles, might also be possible for limited numbers of (adult) deaf speakers.

REFERENCES

[Atk73] Atkinson, J.E. (1973) Aspects of Intonation in Speech: Implications from an Experimental Study of Fundamental Frequency. Ph.D. thesis, Univ. of Connecticut.

[AngKop64] Angelocci, A.A., Kopp, G.A. and Holbrook, A. (1964) The vowel formants of deaf and normal-hearing eleven- to fourteen-year-old boys. J. Speech and Hearing Disorders, 29, 156-170.

[BBN75] Proposal for research on "Laryngeal Control in the Speech of Deaf Children: Assessment of Problems and Implications for Diagnosis and Training of Speech". Bolt, Beranek and Newman, Inc., January, 1975.

[BBN77] Speech training units. Bolt, Beranek and Newman, Inc. Progress Report, December, 1977.

[Ben73] Benade, A. (1973) The physics of brasses. Scientific American, July, 24-35.

[Ber55] van den Berg, J.W. (1955) Uber die Koppelung bei der Stimmenbildung. Z. Phonetik und allgem. Sprachwissenschaft, 8, 281-293 (cited in [Sun75]).

[BerRol78] Bernstein, J., Rollins, A.M. and Stevens, K.N. (1978) Word and syllable concatenation in the speech of the deaf. Presented at a meeting of the A.G. Bell Association for the Deaf, St. Louis, Missouri, June 25, 1978.

[Boo66] Boone, D.R. (1966) Modification of the voices of deaf children. The Volta Review, 48, 686-692.

[Boo80] Boothroyd, A. (1980) Evaluation of speech. The Clarke School for the Deaf, SARP No. 34.

[Cal61] Calvert, D. (1961) Some acoustic characteristics of the speech of profoundly deaf individuals. Ph.D. Thesis, Stanford University, Stanford, CA.

[EguHir69] Eguchi, S. and Hirsh, I.J. (1969) Development of speech sounds in children. Acta Oto-Laryngologica, Suppl. 257, 1-51.

[Eng62] Engelberg, M. (1962) Correction of falsetto voice in a deaf adult. JSHD, 27, 162-164.

[EriLib77] Erickson, D., Liberman, M. and Niimi, S. (1977) The geniohyoid and the role of the strap muscles. Haskins Lab. Progress Rpt., SR-49, 103-110.

[Ewa75] Ewan, W.G. (1975) Explaining the intrinsic pitch of vowels. San Jose Papers in Linguistics, 1, 56-65.

[Ewa79a] Ewan, W.G. (1979) Can intrinsic F0 be explained by source/tract coupling? JASA, 66, 358-362.

[Ewa79b] Ewan, W.G. (1979) Laryngeal behavior in speech. Rpt. of the Phonology Lab, 3, Univ. of Calif., Berkeley.

[Fan56] Fant, G. (1956) On the predictability of formant levels and spectrum envelopes from formant frequencies. In For Roman Jakobsen, M. Halle, H. Lunt and H. MacLean (eds.) The Hague: Mouton.

[Fan60] Fant, G. (1960) Acoustic Theory of Speech Production. The Hague: Mouton.

[Fan72] Fant, G. (1972) Vocal tract wall effects, losses and resonance bandwidths. STL-QPSR, 2-3/1972, 28-52.

[Fis67] Fischer-Jorgensen, E. (1967) Phonetic analysis of breathy (murmured) vowels in Gujarati. Annual Rpt. Inst. of Phonetics, Univ. of Copenhagen, 2, 35-85.

[Fla65] Flanagan, J.L. (1965) Speech Analysis, Synthesis and Perception. New York: Springer-Verlag.

[FlaLan68] Flanagan, J.L. and Landgraf, L.L. (1968) Self-oscillating source for vocal-tract synthesizers. IEEE Trans. on Audio and Electroacoustics, AU-16, 57-64.

[FujLin71] Fujimura, O. and Lindqvist, J. (1971) Sweep-tone measurements of vocal-tract characteristics. JASA, 49, 541-558.

[Gau77] Gauffin, J. (1977) Mechanisms of larynx tube constriction. Phonetica, 34, 307-309.

[Gol76] Goldstein, U.G. (1976) Speaker-identifying features based on formant tracks. JASA, 59, 176-182.

[Gol79] Goldstein, U.G. (1979) LINPC User Documentation. Speech Communication Research Group, Massachusetts Institute of Technology, Cambridge, MA.

[Gol80] Goldstein, U.G. (1980) An articulatory model for the vocal tracts of growing children. Sc.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA.

[Har76] Hardcastle, W.J. (1976) Physiology of Speech Production. New York: Academic Press.

[Hen76] Henke, W.L. (1976) FPRD User Documentation. Speech Communication Research Group, Massachusetts Institute of Technology, Cambridge, MA.

[HolPau69] Hollien, H. and Paul, P. (1969) A second evaluation of the speaking fundamental frequency characteristics of post-adolescent girls. Language and Speech, 12, 119-124.

[Hon81] Honda, K. (1981) Relationship between pitch control and vowel articulation. Presented at the "Vocal-Fold Physiology Conference", Univ. of Wisconsin, May 31-June 4, 1981.

[HonBae81] Honda, K., Baer, T., Hirose, H. and Sawashima, M. (1981) Relationship between vowel articulation and pitch control. JASA, 69, S67 (A).

[HorBis72] Horii, Y. and Bishop, M.E. (1972) Fundamental frequency characteristics of young deaf adults during oral reading. JASA, 52, 146, (A).

[Hor77] Horwich, E.J. (1977) An Examination of the Vowels of Deaf Girls. S.B. Thesis, Dept. of Electrical Engineering and Computer Science, Massachusetts Institute of Technology.

[Hou59] House, A.S. (1959) A note on optimal vocal frequency. JSHR, 2, 55-60.

[HouFai53] House, A.S. and Fairbanks, G. (1953) The influence of consonant environment upon the secondary acoustical characteristics of vowels, JASA, 25, 105-113.

[HouSte56] House, A.S. and Stevens, K.S. (1956) Analog studies of the nasalization of vowels. JSHD, 21, 218-232.

[Hud37] Hudgins, C.V. (1937) Voice production and breath control in the speech of the deaf. Am. Ann. of the Deaf, 82, 338-363.

[IshFla72] Ishizaka, K. and Flanagan, J.L. (1972) Synthesis of voiced sounds from a two-mass model of the vocal cords. Bell Sys. Tech J., 51, 1233-1268.

[IshMat68] Ishizaka, K. and Matsudaira, M. (1968) Analysis of the vibration of the vocal cords. J. Acoust. Soc. Japan, 24, 311-314.

[JanSun74] Jansson, E. and Sundin, H. (1974) A pilot study on coupling between top plate and air volume vibrations. Catgut Acoust. Soc. Newsletter, May 1, 11-15 (cited in [Sun75]).

[KagTre37] Kagen, B. and Trendelenburg, W. (1937) Zum kenntnis der wirkung von kunstlichen ansatzrohren auf die stimmschwingungen. Arch. f. Sprach- und Stemmheilkunde, 1, 129-150 (cited in [Sun75]).

[Kah75] Kahane, J.C. (1975) The developmental anatomy of the human prepubertal and pubertal larynx. Ph.D. Thesis, University of Pittsburgh, Pittsburgh, PA.

[KakHir81] Kakita, Y., Hirano, M. and Ohmaru, K. (1981) Physical properties of the vocal fold tissue: Measurements on excised larynges. In Vocal Fold Physiology, K.N. Stevens and M. Hirano (eds.) Tokyo: Univ. of Tokyo Press.

[Ken76] Kent, R.D. (1976) Anatomical and neuromuscular maturation of the speech mechanism: Evidence from acoustic studies. JSHR, 19, 421-447.

[Lad73] Ladefoged, P. (1973) The features of the larynx. J. Phonetics, 1, 73-83.

[Lea73] Lea, W.A. (1973) Segmental and suprasegmental influences on fundamental frequency contours. In Consonant Types and Tone, Southern California Occasional Papers in Linguistics, 1, L.M. Hyman (ed.).

[Leh70] Lehiste, I. (1970) Suprasegmentals. Cambridge, MA: The M.I.T. Press.

[LehPet61] Lehiste, I. and Peterson, G.E. (1961) Some basic considerations in the analysis of intonation. JASA, 33, 419-425.

[LevSmi74] Levitt, H., Smith, C.R. and Stromberg, H. (1974) Acoustic, articulatory and perceptual characteristics of the speech of deaf children. Presented at the Speech Communication Seminar, Stockholm, Aug. 1-3.

[Lie70] Lieberman, P. (1970) A study of prosodic features. Haskins Lab. Progress Rpt., SR-23, 179-208.

[Lin76] Ling, D. (1976) Speech and the Hearing-Impaired Child: Theory and Practice. Washington, D.C.: A.G. Bell Association for the Deaf.

[Lin-Gau72] Lindqvist-Gauffin, J. (1972) A descriptive model of laryngeal articulation in speech. STL-QPSR, 2-3/1972, 1-9.

[Mag72] Magner, M.E. (1972) A Speech Intelligility Test for Deaf Children. The Clarke School for the Deaf, Northampton, MA.

[Mar68] Martony, J. (1968) On the correction of the voice pitch level for severely hard-of-hearing subjects. Amer. Annals of the Deaf, 113, 195-202.

[Mar77] Martony, J. (1977) Some aspects on speech errors in deaf children. Presented at the Research Conference on Speech-Processing Aids for the Deaf, Gallaudet College, Washington, D.C., May 24-26.

[McGOsb78] McGarr, N.S. and Osberger, M.J. (1978) Pitch deviancy and intelligibility of deaf speech. J. Communication Disorders, 11, 237-247.

[Mea80] Meadow, K.P. (1980) Deafness and Child Development. Berkeley/Los Angeles: The University of California Press.

[Mon76] Monsen, R.B. (1976) Normal and reduced phonological space: the production of English vowels by deaf adolescents. J. Phonetics, 4, 189-198.

[Mon78] Monsen, R.B. (1978) Toward measuring how well hearing-impaired children speak. J. Speech and Hearing Research, 21, 197-219. 1680-1690.

[MorIng68] Morse, P.M. and Ingard, K.U. (1968) Theoretical Acoustics. New York: McGraw-Hill.

[Nic75] Nickerson, R.S. (1975) Characteristics of the speech of deaf persons. The Volta Review, 77, 342-362.

[NicSte77] Unpublished report on Bolt, Beranek and Newman Speech-Training Units.

[NicSte79a] Nickerson, R.S., Stevens, K.N., and Rollins, A.M. (1979) Computers and Sensory Aids. Presented at the Conference on the Speech of the Hearing-Impaired: Research, Training and Personnel Preparation, City University of New York, New York, NY, Oct. 31-Nov. 2.

[NicSte79b] Nickerson, R.S., Stevens, K.N. and Rollins, A.M. (1979) Research on Computer Based Speech Diagnosis and Speech Training Aids for the Deaf. Technical Report No. 4029, Bolt, Beranek and Newman, Inc., Cambridge, MA.

[Oha73] Ohala, J.J. (1973) Explanations for the intrinsic pitch of vowels. Monthly Internal Memorandum, Phonology Lab., Univ. of Calif., Berkeley, 9-26.

[Oha77] Ohala, J.J. (1977) Speculations on pitch regulation. Phonetica, *34*, 310-312.

[OhaEuk78] Ohala, J.J. and Eukel, B.W. (1978) Explaining the intrinsic pitch of vowels. Rpt. of the Phonology Lab., Univ. of Calif., Berkeley, *2*, 118-125.

[ParLev78] Parkhurst, B.G. and Levitt, H. (1978) The effect of selected prosodic errors on the intelligibility of deaf speech. J. Communication Disorders, *11*, 249-256.

[Per69] Perkell, J.S. (1969) *Physiology of Speech Production: Implications of a Quantitative Cineradiographic Study*. Research Monograph No. *53*, The M.I.T. Press, Cambridge, MA. Press.

[Pet76] Petersen, N.R. (1976) Intrinsic fundamental frequency of Danish vowels. Ann. Rpt. of the Inst. of Phonetics, Univ. of Copenhagen, *10*, 1-27.

[Pet46] Peterson, G.E. (1946) Influence of voice quality. The Volta Review, *48*, 640-641.

[PetBar52] Peterson, G.E. and Barney, H.L. (1952) Control methods used in a study of the vowels. JASA, *24*, 175-184.

[PetLeh60] Peterson, G.E. and Lehiste, I. (1960) Duration of syllabic nuclei in English. JASA, *32*, 693-703.

[Rot74] Rothenberg, M. (1974) Glottal noise during speech. STL-QPSR, *2-3/1974*, 1-10.

[Sap78] Sapir, S. (1978) Laryngeal motions during changes in the voice fundamental frequency (F0). Presented at the American Speech and Hearing Convention, San Francisco, CA, November, 1978.

[ShaPie79] Shadle, C.H., Pierrehumbert, J.B. and Lieberman, M.Y. (1979) The intrinsic pitch of vowels in sentence context. JASA, 66, S64, (A).

[Ste-pc] Stevens, K.N., personal communication.

[Ste77] Stevens, K.N. (1977) Physics of laryngeal behavior and laryngeal modes. Phonetica, 34, 264-279.

[SteKal75] Stevens, K.N., Kalikow, D.N. and Willemain, T.R., (1975) A miniature accelerometer for detecting glottal waveforms and nasalization. JSHR, 18, 594-599.

[SteNic79] Stevens, K.N., Nickerson, R.S. and Rollins, A.M. (1979) Suprasegmental and postural aspects of speech production and their effect on articulatory skills and intelligibility. Presented at the Conf. on the Speech of the Hearing Impaired: Research, Training and Personnel Preparation, City University of New York, Oct.31-Nov.2.

[StrLev79] Stromberg, H. and Levitt, H. (1979) Multiple linear regression analysis of errors in deaf speech. In Speech Communication Papers Presented at the 97th Meeting of the Acoustical Society of America, J.J. Wolf and D.H. Klatt (eds.) New York: Acoustical Society of America.

[Sun75] Sundberg, J. (1975) Formants and fundamental frequency control in singing. An experimental study of coupling vocal tract and voice source. STL-QPSR, 1/1975, 65-78.

[SunGau78] Sundberg, J. and Gauffin, J. (1978) Waveform and spectrum of the glottal voice source. STL-QPSR, 2,3/1978, 35-50.

[Zee78] Zee, E. (1978) The interaction of tone and vowel quality. UCLA Working Papers in Phonetics, 41, 53-67.

APPENDIX 1

Table A1.1 - F0(V)

speaker	/i/	/I/	/ɛ/	/æ/	/ɑ/
D1	268	259	200	206	191
D2	277	281	240	231	222
D3	355	343	311	291	288
D4	315	305	274	290	292
D5	302	287	278	268	271
D6	290	274	286	266	256
D7	327	334	326	310	329
D8	345	350	352	379	379
D9	353	298	255	290	226
D10	328	292	256	218	219
D11	285	226	216	231	190
D12	302	288	248	238	202
D13	273	256	231	212	216
D14	255	249	244	244	225
D15	332	307	275	261	261
D16	242	225	209	211	208
D17	302	303	282	298	274
D18	210	208	199	195	194
D19	315	308	302	296	288
D20	363	335	346	331	342
H1	238	246	236	240	250
H2	242	235	221	215	215
H3	205	185	189	184	188
H4	232	223	234	220	216
H5	230	227	216	216	226
H6	198	191	182	178	180
H7	240	249	224	231	230
H8	215	212	204	201	200
H9	223	217	206	214	211
H10	212	206	204	202	211
H11	203	198	190	192	195
H12	223	224	221	224	222
H13	223	227	219	220	222

F0(V)

speaker	/ʌ/	/ɝ/	/ʊ/	/u/	/ɑy/
D1	196	199	214	258	224
D2	250	218	251	298	245
D3	279	301	320	342	318
D4	294	276	312	325	321
D5	265	276	285	288	280
D6	259	265	266	279	260
D7	292	315	321	334	320
D8	374	349	350	344	378
D9	249	294	281	326	260
D10	236	271	301	331	261
D11	206	216	201	256	192
D12	212	239	236	285	254
D13	231	221	236	288	219
D14	249	275	291	290	228
D15	291	270	259	335	290
D16	210	229	229	246	202
D17	286	310	300	331	289
D18	192	224	222	221	198
D19	328	299	316	308	318
D20	356	342	356	321	361
H1	201	230	232	244	220
H2	206	206	231	226	208
H3	186	186	192	200	181
H4	209	224	221	229	212
H5	221	221	228	224	224
H6	184	181	202	194	180
H7	220	221	235	239	229
H8	201	201	209	222	208
H9	211	214	211	216	212
H10	209	204	214	212	198
H11	196	196	204	201	195
H12	221	225	225	221	211
H13	224	222	225	224	222

Table A1.2 - $\frac{F0(V)}{F0(/\text{ɑ}/)}$

speaker	/i/	/I/	/ɛ/	/æ/	/ɑ/
D1	1.403	1.354	1.046	1.078	1.000
D2	1.244	1.263	1.079	1.039	1.000
D3	1.235	1.193	1.083	1.013	1.000
D4	1.077	1.043	.936	.992	1.000
D5	1.112	1.058	1.023	.986	1.000
D6	1.132	1.069	1.117	1.039	1.000
D7	.994	1.016	.992	.943	1.000
D8	.911	.924	.931	1.000	1.000
D9	1.562	1.317	1.127	1.282	1.000
D10	1.501	1.335	1.171	.994	1.000
D11	1.500	1.190	1.138	1.217	1.000
D12	1.490	1.422	1.222	1.173	1.000
D13	1.264	1.184	1.104	.983	1.000
D14	1.133	1.107	1.083	1.083	1.000
D15	1.270	1.175	1.053	1.000	1.000
D16	1.165	1.084	1.006	1.018	1.000
D17	1.020	1.107	1.032	1.087	1.000
D18	1.084	1.074	1.026	1.006	1.000
D19	1.096	1.071	1.052	1.030	1.000
D20	1.061	.978	1.011	.967	1.000
H1	.953	.984	.945	.960	1.000
H2	1.124	1.093	1.029	1.000	1.000
H3	1.093	.987	1.007	.980	1.000
H4	1.071	1.031	1.081	1.029	1.000
H5	1.017	1.003	.956	.956	1.000
H6	1.102	1.061	1.014	.986	1.000
H7	1.044	1.083	.973	1.005	1.000
H8	1.075	1.060	1.019	1.006	1.000
H9	1.057	1.027	.976	1.012	1.000
H10	1.002	.975	.964	.959	1.000
H11	1.043	1.015	.974	.987	1.000
H12	1.004	1.007	.994	1.006	1.000
H13	1.004	1.020	.983	.989	1.000

$$\frac{\mathrm{F0(V)}}{\mathrm{F0(/ɑ/)}}$$

speaker	/ʌ/	/ɜ/	/ʊ/	/u/	/ɑy/
D1	1.026	1.039	1.118	1.346	1.170
D2	1.124	.978	1.129	1.337	1.101
D3	.970	1.048	1.113	1.191	1.104
D4	1.004	.994	1.068	1.111	1.098
D5	.977	1.018	1.051	1.060	1.032
D6	1.010	1.034	1.039	1.088	1.015
D7	.890	.958	.977	1.015	.973
D8	.987	.921	.924	.908	.997
D9	1.099	1.298	1.243	1.442	1.149
D10	1.080	1.240	1.377	1.514	1.194
D11	1.086	1.138	1.059	1.349	1.013
D12	1.049	1.179	1.167	1.409	1.253
D13	1.069	1.023	1.092	1.330	1.012
D14	1.106	1.222	1.294	1.289	1.011
D15	1.115	1.034	.990	1.282	1.110
D16	1.012	1.102	1.102	1.187	.976
D17	1.046	1.132	1.096	1.210	1.055
D18	.994	1.155	1.148	1.142	1.019
D19	1.139	1.039	1.100	1.070	1.104
D20	1.040	1.000	1.040	.938	1.055
H1	.805	.920	.930	.975	.880
H2	.959	.959	1.076	1.052	.965
H3	.993	.993	1.027	1.067	.967
H4	.965	1.035	1.023	1.058	.983
H5	.978	.978	1.006	.989	.989
H6	1.021	1.007	1.125	1.076	1.000
H7	.956	.962	1.022	1.038	.995
H8	1.006	1.006	1.044	1.112	1.038
H9	1.000	1.012	1.000	1.024	1.006
H10	.988	.964	1.012	1.006	.935
H11	1.006	1.006	1.045	1.032	1.000
H12	.994	1.011	1.011	.994	.949
H13	1.006	1.000	1.011	1.006	1.000

Table A1.3 - F1(V)

speaker	/i/	/I/	/ɛ/	/æ/	/ɑ/
D1	467	428	774	835	978
D2	458	359	800	902	908
D3	447	532	888	1005	1052
D4	457	635	781	962	940
D5	595	571	700	899	899
D6	537	498	534	498	629
D7	535	565	630	869	949
D8	472	429	479	630	760
D9	552	681	724	722	834
D10	440	533	629	805	784
D11	642	752	868	888	818
D12	647	555	713	805	746
D13	360	442	758	795	838
D14	558	515	758	849	878
D15	407	419	648	672	808
D16	345	508	641	670	729
D17	762	715	901	859	858
D18	438	402	675	570	808
D19	522	550	638	757	819
D20	550	370	546	566	661
H1	380	522	720	832	834
H2	400	510	676	790	728
H3	375	457	708	811	644
H4	458	583	688	725	674
H5	377	546	680	878	790
H6	382	533	704	977	779
H7	287	503	681	715	671
H8	382	444	631	782	742
H9	395	550	674	682	696
H10	412	466	706	862	925
H11	385	476	630	705	701
H12	378	412	655	648	736
H13	420	445	635	678	671

F1(V)

speaker	/ʌ/	/ɝ/	/ʊ/	/u/	/ɑy/
D1	973	650	741	325	726
D2	905	445	524	310	652
D3	---	636	721	350	844
D4	880	701	635	363	698
D5	700	600	609	529	664
D6	494	505	520	518	647
D7	805	482	636	401	894
D8	695	361	358	436	690
D9	869	718	639	471	869
D10	771	554	620	371	788
D11	826	801	761	673	868
D12	830	622	670	506	824
D13	798	544	588	311	519
D14	903	478	404	335	962
D15	679	598	575	415	520
D16	679	524	524	392	685
D17	846	746	752	645	852
D18	893	444	481	422	460
D19	854	518	585	320	720
D20	664	420	512	---	631
H1	790	592	550	421	810
H2	632	504	550	435	585
H3	590	482	529	380	514
H4	655	580	665	459	601
H5	772	691	636	398	775
H6	865	518	567	376	830
H7	682	469	489	414	679
H8	621	508	486	392	718
H9	675	612	616	424	681
H10	826	528	495	422	839
H11	638	500	508	392	622
H12	736	456	501	372	684
H13	685	490	522	450	602

Table A1.4 - F2(V)

speaker	/i/	/I/	/ɛ/	/æ/	/ɑ/
D1	1910	2078	1870	1920	1480
D2	2440	2518	1961	1934	1609
D3	2695	2605	2098	2130	1551
D4	2338	2114	1855	1866	1629
D5	1885	1930	1979	1904	1800
D6	2050	2092	2184	2010	1801
D7	2235	2286	2042	2000	1769
D8	2068	1940	2181	2194	1542
D9	2325	1966	1702	1899	1359
D10	2138	2041	1854	1778	1471
D11	2030	1675	1765	1712	1374
D12	1652	1937	1838	1664	1522
D13	2585	2208	2008	1965	1238
D14	2220	2006	1767	1878	1448
D15	1967	1848	1599	1484	1306
D16	2518	2062	1859	2000	1176
D17	1827	1834	1650	1741	1479
D18	2308	2455	1945	1952	1184
D19	2050	2075	1850	1857	1686
D20	1817	2184	1735	1852	1792
H1	2552	2195	2146	2119	1291
H2	2593	2249	2122	2170	1032
H3	2952	2259	2118	2210	1125
H4	3010	2392	2199	2338	1279
H5	2885	2410	2279	2264	1240
H6	2713	2198	1891	1803	1335
H7	2530	2058	1795	1949	1166
H8	2433	2228	1989	1925	1230
H9	2417	2127	1918	1938	1281
H10	2607	2319	2094	2007	1294
H11	2337	2042	1835	2046	1238
H12	2500	2254	1945	2080	1255
H13	2425	2075	1900	1788	1238

F2(V)

speaker	/ʌ/	/ɝ/	/ʊ/	/u/	/ɑy/
D1	1688	1676	1526	1375	1858
D2	1622	1744	1666	1959	2176
D3	----	1508	1266	1055	2240
D4	1781	1952	1569	1500	2188
D5	1620	1546	1668	1774	1838
D6	1919	2050	1945	1971	1988
D7	1647	1819	1698	1261	1941
D8	1906	1956	2238	1859	1916
D9	1484	1676	1145	1410	1576
D10	1559	1689	1699	1670	1809
D11	1430	1579	1412	1412	1628
D12	1552	1564	1499	1419	1608
D13	1290	1879	1407	1092	2181
D14	1732	1610	1349	1098	1569
D15	1308	1364	1394	1324	1686
D16	1135	1572	1176	1135	1740
D17	1504	1469	1417	1494	1747
D18	1288	1756	1366	1692	2113
D19	1825	2001	1818	1818	1868
D20	1739	2062	1305	----	1810
H1	1438	1521	1175	1390	1681
H2	1368	1519	1251	1446	1987
H3	1530	1550	1266	1556	1880
H4	1510	1725	1270	1391	2065
H5	1589	1760	1318	1312	2061
H6	1610	1709	1423	1432	1647
H7	1402	1456	1201	1330	1619
H8	1469	1575	1180	1297	1773
H9	1616	1639	1368	1399	1854
H10	1606	1605	1288	1446	1750
H11	1476	1491	1281	1518	1672
H12	1559	1685	1179	1440	1755
H13	1545	1563	1241	1370	1828

APPENDIX 2

Table A2.1

F0 Data for Nasal Consonants in Syllable-Initial Position

Deaf Speakers (n = 20)

target noun	number of tokens	F0 (Nasal)	F0 (Vowel)
mitt	16	276	294
man	17	260	262
mug	17	260	261
moon	19	287	291
net	13	276	276
nook	15	275	267

Hearing Speakers (n = 13)

target noun	number of tokens	F0 (Nasal)	F0 (Vowel)
mitt	13	216	218
man	13	216	212
mug	13	216	208
moon	13	218	216
net	13	210	199
nook	13	216	215

Table A2.2

F0 Data for Nasal Consonants in Syllable-Final Position

Deaf Speakers (n = 20)

target noun	number of tokens	F0 (Vowel)	F0 (Nasal)
time	18	253	235
man	18	261	240
dawn	15	240	229
burn	16	265	242
moon	20	293	251
king	7	302	260

Hearing Speakers (n = 13)

time	10	214	215
man	12	210	202
dawn	12	208	202
burn	11	213	211
moon	12	216	199
king	12	222	205

Table A2.3

Grouped F0 Data for Syllable-Initial /m/ Followed by High and Low Vowels

High Vowels /I,u/

group	number of tokens	F0 (Consonant)	F0 (Vowel)
hearing	26	217	217
A-1	13	300	309
A-2	13	286	294
B	4	219	225
C	5	273	299

Low Vowels /æ,ʌ/

hearing	26	216	210
A-1	10	292	310
A-2	12	258	242
B	4	214	221
C	8	248	251

Table A2.4

F0 Data for /r,l,w/ in Syllable-Initial Position

Deaf Speakers (n = 20)

target noun	number of tokens	F0 (Consonant)	F0 (Vowel)
rag	16	252	252
rug	15	251	244
root	16	265	296
leak	17	270	297
lid	16	264	276
light	14	249	270
well	13	251	272
wall	14	238	237
word	13	255	259
wool	13	259	273

Table A2.5

F0 Data for /r,l,w/ in Syllable-Initial Position

Hearing Speakers (n = 13)

target noun	number of tokens	F0 (Consonant)	F0 (Vowel)
rag	13	210	205
rug	13	213	203
root	12	209	218
leak	13	217	220
lid	13	215	206
light	13	214	204
well	12	210	209
wall	13	212	212
word	12	212	204
wool	13	206	207

Table A2.6

F0 Data for /l/ in Syllable-Final Position

Deaf Speakers (n = 20)

target noun	number of tokens	F0 (Vowel)	F0 (/l/)
well	15	264	222
wall	17	242	212
wool	19	277	222

Hearing Speakers (n = 13)

well	8	206	202
wall	7	209	217
wool	6	203	203

Table A2.7

Grouped F0 Data for Syllable-Initial /r,l,w/ Followed by High and Low Vowels

High Vowels /i,I,u/

group	number of tokens	F0 (Consonant)	F0 (Vowel)
hearing	38	214	214
A-1	13	285	315
A-2	20	269	288
B	6	215	228
C	10	268	296

Low Vowels /æ,ɑ,ʌ/

group	number of tokens	F0 (Consonant)	F0 (Vowel)
hearing	39	212	207
A-1	11	276	290
A-2	20	254	238
B	6	186	198
C	8	235	234

Table A2.8

F0 Data for Vowels Preceded by Voiced Labial Stops

Deaf Speakers (n = 20)

target noun	number of tokens	F0 (10 msec)	F0 (20 msec)	F0 (center)
beak	14	310	311	308
bid	13	288	290	280
bed	13	270	272	256
bag	14	274	276	267
box	16	261	262	257
bug	15	281	283	271
bird	15	284	283	273
book	14	300	300	286
boot	13	305	310	303
bike	14	288	288	285

Table A2.9

F0 Data for Vowels Preceded by Voiced Labial Stops

Hearing Speakers (n = 13)

target noun	number of tokens	F0 (10 msec)	F0 (20 msec)	F0 (center)
beak	13	238	238	226
bid	12	233	231	217
bed	13	225	220	211
bag	13	219	218	206
box	13	218	219	210
bug	12	224	220	200
bird	13	218	219	206
book	13	223	223	215
boot	13	233	229	217
bike	12	216	215	204

Table A2.10

Grouped F0 Data for High and Low Vowels Preceded by Voiced Labial Stops

High Vowels /i,I,u/

group	number of tokens	F0 (10 msec)	F0 (20 msec)	F0 (center)
hearing	38	235	232	220
A-1	15	314	316	316
A-2	15	301	305	290
B	4	251	244	238
C	6	303	309	308

Low Vowels /æ,ɑ,^/

group	number of tokens	F0 (10 msec)	F0 (20 msec)	F0 (center)
hearing	38	220	219	206
A-1	14	312	315	319
A-2	17	273	276	257
B	5	207	204	199
C	9	242	242	230

Table A2.11

F0 Data for Vowels Preceded by Voiceless Stops

Deaf Speakers (n = 20)

target noun	number of tokens	F0 (10 msec)	F0 (20 msec)	F0 (center)
peak	16	296	299	287
pit	17	301	297	289
pet	18	284	290	277
pipe	18	290	292	276
top	18	282	284	261
toot	15	309	310	299
time	13	268	267	245
cat	16	283	289	261
cut	14	292	287	264
cook	17	304	309	293
king	18	306	307	290

Table A2.12

F0 Data for Vowels Preceded by Voiceless Stops

Hearing Speakers (n = 13)

target noun	number of tokens	F0 (10 msec)	F0 (20 msec)	F0 (center)
peak	13	241	234	220
pit	13	249	249	234
pet	13	243	236	225
pipe	13	233	231	212
top	13	234	232	222
toot	13	244	242	225
time	13	230	230	212
cat	13	238	237	220
cut	13	229	224	215
cook	13	255	248	234
king	13	244	242	220

Table A2.13

Grouped F0 Data for High and Low Vowels Preceded by Voiceless Stops

High Vowels /i,I,u/

group	number of tokens	F0 (10 msec)	F0 (20 msec)	F0 (center)
hearing	39	245	242	226
A-1	12	312	316	309
A-2	18	317	318	306
B	6	238	233	219
C	12	300	298	288

Low Vowels /æ,ɑ,ʌ/

hearing	39	234	231	219
A-1	13	320	324	311
A-2	19	305	308	274
B	6	218	218	194
C	10	243	238	216

Table A2.14

Grouped F0 Data for Vowels Preceded by the Voiceless Stops /p,t,k/

group	number of tokens	F0 (10 msec)	F0 (20 msec)	F0 (center)
		/p/		
hearing	52	241	238	223
A-1	20	312	318	315
A-2	25	305	307	289
B	8	229	229	213
C	16	281	278	263
		/t/		
hearing	39	236	235	220
A-1	12	318	320	309
A-2	19	295	297	272
B	6	231	231	207
C	9	265	263	249
		/k/		
hearing	52	241	238	222
A-1	18	322	326	315
A-2	26	315	318	289
B	7	225	227	207
C	14	266	262	245