ARIPUC 21, 1987, p. 29-43

PERCEPTUAL COMPENSATION FOR SEGMENTALLY CONDITIONED FUNDAMENTAL FREQUENCY PERTURBATION*

NIELS REINHOLT PETERSEN

The paper deals with the role of two types of segmentally conditioned fundamental frequency (Fo) perturbation in speech perception, i.e. the differences in Fo between high and low vowels and the effect of prevocalic obstruent voicing on Fo. The literature relevant to this issue is reviewed, and the results of a supplementary experiment on Fo as a cue to vowel identity are reported. On this basis three major points are made: (1) Although the pitch of vowels is influenced psychoacoustically by their spectral shape, this effect is too small to be of any significant importance in speech perception. (2) The two types of Fo perturbation are not treated differently in the speech perception process, and (3) irrespective of type, the segmentally conditioned Fo variation can function as a cue both to segment identity and to the identity of prosodic categories, depending on the actual demands upon the speech perception system.

I. INTRODUCTION

It is a well-established fact that the fundamental frequency (Fo) of speech is not only determined by higher-level linguistic information such as sentence type, stress pattern, and tone but also by the segments constituting the utterance. The latter, microprosodic Fo variation can be described as having two sources, one being the inherent Fo differences between segments and the other being coarticulatory effects of segments (consonants in particular) on Fo in adjacent ones. Although the underlying mechanisms are still subject to discussion, both types of micro-

---

*) *Also in* Phonetica *43/1-3, 1986, p. 31-42.*

prosodic variation are generally assumed to be a result of physiological and/or acoustic constraints of the speech production apparatus and thus beyond the active control of the speaker (see e.g. Ohala 1973, Reinholt Petersen 1978, 1983, Silverman 1984a).

The present paper will deal with the role of microprosodic phenomena in the perception of speech, concentrating upon intrinsic Fo of vowels (i.e. the tendency for high vowels to have a higher Fo than low vowels, ceteris paribus) and the influence of obstruent voicing on Fo in the following vowel (Fo being lower after voiced than after voiceless obstruents) as examples of the two types of microprosodic Fo variation. In connection with the discussion of phonetic explanations for the development (or nondevelopment) of linguistic tones from segmental effects on Fo (Hombert 1976a,b, 1978, Ohala 1978, 1982), the question was raised whether the two types of segmentally conditioned Fo perturbation are treated differently by the perceptual system. Considering the tendency for the development of tones from initial consonant voicing to be far more frequent than the development of tones from vowel height (in fact, no unambiguous cases appear to have been reported), it has been assumed that consonantal influences on Fo are detected by listeners, whereas intrinsic differences among vowels go unnoticed, presumably due to some sort of perceptual compensatory effect (Hombert 1976b, Silverman 1984a). Since this issue is relevant not only to theories of tonogenesis but also, of course, of primary interest to general phonetic theory, the following study will attempt to give a comprehensive review of experimental data available in the literature that may illuminate the question whether intrinsic Fo differences and consonantal influences on Fo are treated differently in terms of compensation in the process of speech perception. In addition it is the aim to report the results of an experiment carried out on a point which - as will appear - has been neglected in previous research.

II. ACOUSTIC CHARACTERISTICS

Before turning to the data on the perception of microprosodic Fo variation it may be appropriate to survey briefly the acoustic characteristics of the two types of segmental influence considered here. The effect of vowel height on Fo can be adequately described as a shift parallelling the underlying Fo contour, i.e. the contour determined by higher-level linguistic information, perhaps with a tendency for the effect to be smaller at the beginning and end of the vowel than in the middle. In stressed vowels the magnitude of the difference between high and low vowels ranges between about 4 and 20%, with typical values of approximately 10%; in unstressed vowels the differences tend to be somewhat smaller.

With respect to the influence of obstruent voicing on Fo in the following vowel there seems to be general agreement that, at vowel onset, Fo is lower after a voiced than after a voiceless obstruent, and that the difference - although being gradually reduced - often persists during the greater part of the vowel. There is less consensus, however, about the direction of the Fo contour following the onset of the vowel. Mohr (1971), Lea (1973), and Hombert (1978) report a rising contour after voiced and a falling contour after a voiceless obstruent, whereas more recent research (Haggard et al. 1981, Ohde 1984, 1985, Silverman 1984b) has shown that Fo may have a falling contour after both obstruent types, but more steeply falling after voiceless than after voiced obstruents. Obviously, as is the case for intrinsic Fo of vowels, the effect of obstruent voicing on Fo comes out as a deviation from the underlying Fo contour and can be described as a fall or rise <u>in relation to that contour</u>. This point is illustrated in figure 1, which shows Fo contours in the vowel [i] after [p] and [b] under three different tonal conditions. The tracings are derived from unpublished material on Dogri, an Indo-Aryan language spoken in the Jammu and Kashmir State in India.

The magnitude of the differences reported in the literature seems to vary a great deal, partly due to the use of different experimental methodologies, and partly due to the fact that speakers display a considerable amount of variation even under identical experimental conditions, but roughly estimated the differences in stressed vowels are typically in the neighborhood of 10-15% in the early part of the vowel, being reduced proportionally towards the end, where it is in most cases eliminated.

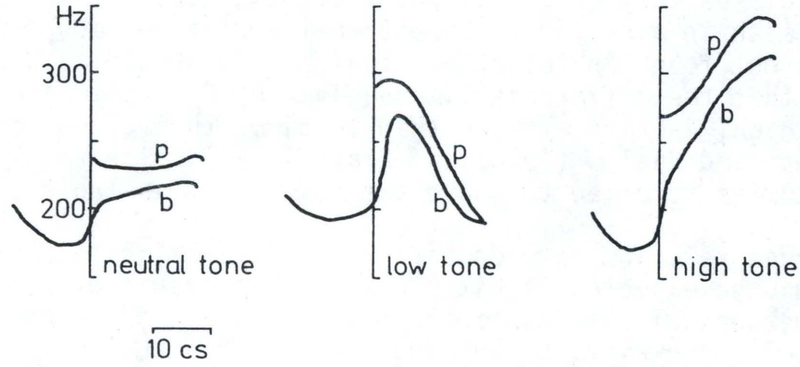

Figure 1

Fo tracings (average of 5 Fo curves) in the vowel [i] after [b] and [p] under three tonal conditions in Dogri. The tracings are lined up around the onset of the vowel.

III. PSYCHOACOUSTIC CONSIDERATIONS

In terms of detectability there seems to be no essential difference between the two types of segmental Fo perturbation described above. Both types exceed the difference limen of about 0.25% for Fo in monotone synthetic steady-state vowels obtained by Flanagan and Saslow (1958) and Klatt (1973); and, what is more important, they also exceed the considerably greater difference limen (2.1%) obtained by Klatt in vowels with linearly changing Fo.

There is, however, evidence from a number of studies that the intrinsic Fo differences among vowels may be, to some extent, reduced in the auditory process as a result of an interaction between vowel quality and pitch. This was first reported by Hombert (1976b), who found that in pairs of high and low vowels the low vowel was judged to be higher in pitch than the high vowel when the two vowels had the same Fo. From this follows that, in order for the two vowels in a pair to be judged equal in pitch, the low vowel must be lower than the high vowel in Fo. Hombert does not state how much lower, but since listeners gave 90% or more correct 'different' judgements when the two vowels in a pair differed by 5 Hz in Fo, he estimates the Fo distance at equal pitch not to exceed 5 Hz. A more exact maximum estimate can, in fact, be arrived at from the distribution of responses between equal Fo and a 5-Hz Fo difference, and under the assumption of normal ogive psychometric functions. On this basis the maximum Fo difference at equal pitch is estimated to 1.26 Hz in i-a pairs and 1.32 Hz in u-a pairs. More recently the effect of vowel quality on pitch has been investigated by Chuang and Wang (1978), Stoll (1984) and Di Cristo (1985). The two former studies find, in agreement with Hombert (1976b), that at equal pitch low vowels have to be lower in Fo than high vowels. Di Cristo (1985) found no clear effect; the reported biases vary greatly among subtests, not only in magnitude but also in direction. Directional variation among subtests was not found in the other studies. It should be added, perhaps, that the difference limens given by Di Cristo are remarkably high, 4-6%, i.e. more than 10 times the values found by Flanagan and Saslow (1958) and Klatt (1973). The results of the studies referred to above are summarized in table I.

The observed relation between pitch and vowel height is indeed what should be expected if pitch bias were a result of some sort of perceptual compensation for intrinsic Fo differences among vowels. However, Chuang and Wang (1978) and Stoll (1984) present evidence indicating that the pitch biases observed for vowels can be accounted for by pitch perception processes common to the general class of complex tone signals, of which speech vowels are, of course, a subset. The results of Stoll (1984) are particularly interesting in this respect. In this study, which examined not only high and low vowels but the entire set of German tense vowels plus schwa, Stoll found a very close correlation between experimentally obtained pitch biases and biases predicted by a pitch extraction algorithm devised by Terhardt and co-workers, which, on the basis of

Table I

Pitch biases in percent of the reference

	i-a	u-a	y-a
Hombert (1976b)	1.26	1.32	
Chuang and Wang (1978)	1.53	2.80	
Stoll (1984)	0.96 0.71	0.78 0.89	1.42
Di Cristo (1985)	1.30 -0.80 0.00 -0.65 2.41		

Note that in the case of Hombert (1976b) the values listed are estimated maximum values. Positive values indicate $Fo_{high} > Fo_{low}$*, negative values* $Fo_{high} < Fo_{low}$ *at equal pitch.*

our present knowledge of pitch perception, describes the relation between spectral properties and pitch of complex tones (see Terhardt 1979, Terhardt et al. 1982a,b).

Thus, the intrinsic Fo differences among vowels and the Fo differences due to consonant voicing can be said to be treated differently by the auditory system. The former type of Fo perturbation is reduced as a result of the interaction between spectral shape and pitch, while the latter is, of course, left unaffected. But it should be pointed out that the pitch bias effect observed in the above studies is quite small, little more than 1% on the average and not exceeding 3%, even if the experimental conditions - direct comparison of monotone stimuli - may be assumed to enhance such an effect as compared to more speech-like conditions. Since the Fo differences between high and low vowels are typically about 10%, the pitch bias effect, however interesting it is from a psychoacoustic point of view, can hardly be assumed to play anything but a marginal role in the perception of speech.

IV. Fo PERTURBATION AND LINGUISTIC CATEGORIES

The fact that the pitch bias effect cannot be assumed to compensate for Fo differences among vowels does not preclude, of course, that perceptual compensation can take place at a higher (i.e. linguistic) level where it is conditioned not by the spectral properties as such but by the phonetic characteristics of the vowels. However, if such a higher-level compensatory effect may be assumed to exist, there is no obvious reason why it should be restricted to operate on Fo differences among vowels, and not on the perturbations caused by the voicing state of the preceding obstruent. It is true, as argued by Hombert et al. (1976), that the effect of vowel height may be more closely associated with the conditioning factor in perception because it occurs simultaneously with it, compared with the effects of obstruent voicing, where the perturbation and the factor causing it are located in different segments. On the other hand, the effect of obstruent voicing on Fo in the following vowel is analogous to the influence of consonant place of articulation on the formant frequencies of adjacent vowels. This effect has been shown to be adjusted for perceptually in that the perceptual boundaries between vowel phonemes are shifted systematically as a function of the place of articulation of the neighboring consonants (Lindblom and Studdert-Kennedy 1967). Thus, it can reasonably be argued that perceptual compensation for perturbatory effects within one segment can operate whether the conditioning factor occurs in that segment or in neighboring segments. Consequently, there seems to be no firm basis, theoretically, for assuming that the two types of segmental Fo perturbation dealt with here are treated differently in the speech perception process; both types may or may not be compensated for.

The question whether or not segmentally conditioned Fo perturbations are actually compensated for can be addressed in terms of the manner in which they interfere with the perception of the identity of the segments themselves on one side, and with the perception of prosodic categories on the other. If a segmental effect on Fo is not compensated for, i.e. it is heard by the listeners, Fo could, at the segmental level, be able to function as a perceptual cue to the phonetic feature which is responsible for that effect and, at the prosodic level, the Fo perturbations could be expected to distort the prosodic structure of an utterance as a function of the segmental composition of the utterance. Conversely, if a segmental effect on Fo is perceptually compensated for, it will not interfere with the prosodic information of the utterance, nor can it be expected to act as a cue to segment identity.

With respect to the influence of Fo on the perception of segments, there is evidence from a number of studies using synthetic speech that Fo in a vowel can be a cue to the voicing category of the preceding obstruent (Haggard et al. 1970, Fujimura 1971, Abramson 1975, Massaro and Cohen 1976, 1977, Abramson and Erickson 1978, Haggard et al. 1981, Bernstein 1983, Abramson and Lisker 1985). In cases where the voice onset time, which is the primary cue to the voicing distinction, is ambiguous, a low Fo in the following vowel will favor 'voiced' responses, whereas a higher Fo will favor 'voiceless' responses. According to Massaro and Cohen (1976, 1978), and Haggard et al. (1981), the most salient Fo characteristic in the vowel for the perception of voicing seems to be the Fo level at vowel onset, rather than the general level in the vowel or the direction (rising versus falling) of the Fo contour following vowel onset. The results of Abramson (1975), Abramson and Erickson (1978), and Abramson and Lisker (1985) suggest, however, that the general Fo level may also be of some importance. In their studies, rising versus falling Fo in the vowel was found to affect the categorization of voicing, and more effectively so when the duration of the Fo movement was 100 and 150 ms instead of 50 ms, i.e. when the 'Fo perturbation' lasted longer. Now, since in natural speech onset Fo, general Fo level, and in some cases also the direction of the initial slope are covarying as a function of consonant voicing, it seems beyond reasonable doubt that the Fo perturbation encountered in natural speech can be used by the perceptual system determining the voicing state of the preceding obstruent.

The question whether intrinsic Fo can contribute to the perception of vowel height has, to my knowledge, not been investigated experimentally. It could be argued that the cases in which Fo can be a significant cue to the identification of vowels are those in which two vowels are close together in vowel height and, consequently, in intrinsic Fo. However, in the Danish material reported by Reinholt Petersen (1978), the Fo differences between adjacent vowels on the vowel height scale were found to be about 5% on the average (ranging from 0 to 11%); and the differences between /i:/ and /e:/ and between /u:/ and /o:/, which are very close in vowel height, averaged 5.5%. Differences of this magnitude may very well be hypothesized to play a role in the perceptual categorization of vowels, particularly if the vowels are closely spaced as is the case in the Danish examples referred to above.

At the prosodic level of speech the results of a study on Danish (Rosenvold 1981) suggest that intrinsic Fo differences between high and low vowels are compensated for in the perception of stress. In Danish, a stressed syllable is characterized by a relatively low Fo, followed by a high Fo in the first posttonic syllable (see, e.g. Thorsen 1984). In an identification experiment, Rosenvold demonstrated that raising the Fo of the second syllable in an isolated two-syllable word from a level equal to that of the first syllable, would shift the perceived stress position from the second to the first syllable, and - what is important here - Fo had to be raised to a higher level for the shift to occur when the vowel in the second syllable was high than when it was low. Not only the direction of the change of boundary but also - and particularly - its magnitude, which was approximately 10%, give strong support to the view that the speech perception system can compensate for intrinsic Fo differences between high and low vowels. And at the same time, the magnitude of the effect rules out the possibility of a psychophysical pitch bias effect to be involved to any significant degree.

Whether similar boundary shifts may occur as a result of perceptual compensation for Fo perturbation due to consonant voicing has been investigated by Abramson and Erickson (1978) in a study on the identification of tonal categories in Thai. Their stimuli were CV syllables synthesized with voice onset times appropriate for [bɑ], [pɑ], and [pʰɑ], and combined with the members of a fan-shaped Fo continuum, in which all Fo variants started from a common origin at 120 Hz and moved towards end points ranging from 92 to 152 Hz. For the sake of clarity the Fo continuum has been redrawn in figure 2. A number of Thai subjects were to identify the stimuli as having low, mid, or high tone. As it turned out, the boundaries between the tonal categories (indicated by arrows in fig. 2) were influenced by the preceding stop consonant, but in a somewhat ambiguous manner.

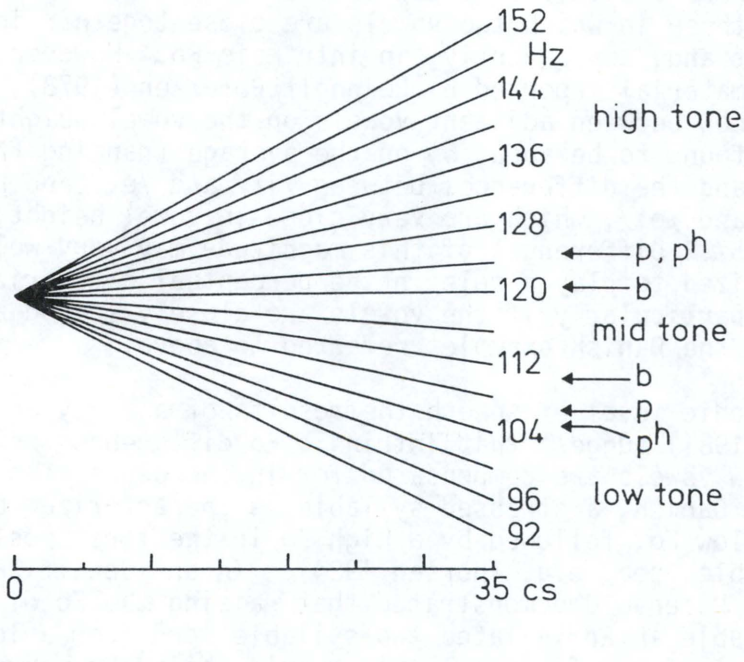

Figure 2

Schematic representation of the Fo variation in the stimuli used by Abramson and Erickson (1978). The boundaries between tonal categories are indicated by arrows (redrawn on the basis of figures in Abramson and Erickson, 1978).

Before dealing further with the results it may be worthwhile to discuss how the tonal boundaries could be expected to vary in the continuum used by Abramson and Erickson, if Fo perturbation due to consonant voicing is perceptually compensated for. This depends, of course, on which aspect of the Fo variation the listeners focus upon in their identification of the tonal categories. In the study at hand the listeners had two (covarying) cues on which to base their judgments, the slope of the Fo contour and the end point (fig. 2).

If the absolute frequency of the end point, irrespective of slope (or, perhaps, the general Fo level, which directly depends on the end point) were the cue to tone identification, then - assuming compensation for the effect of consonant voicing - the boundary between two adjacent tonal categories should be expected at lower end points after voiced than after voiceless stops.

Conversely, if the Fo slope alone were the cue, the boundary between the low and mid tones and between the mid and high tones should be expected at less steeply negative and more steeply positive slopes, respectively, after voiced than after voiceless stops. In other words, the slope representing the boundary between two adjacent tonal categories should be positive after voiced and negative after voiceless stops in relation to a 'neutral' (i.e. unperturbed) boundary slope. And, if this interpretation is expressed in terms of end point values, the boundaries should be expected at higher end points on the stimulus axis after voiced than after voiceless consonants, since a slope with a higher end point is more steeply rising - or less steeply falling - than a slope with a lower end point.

Thus, it is seen that the assumption of perceptual compensation for the effect of consonant voicing will predict opposite directions of boundary shifts depending on whether Fo level or Fo slope is the more important cue to the perception of tonal categories.

The ambiguity of Abramson and Erickson's results rests on the fact that their listeners put the high-mid boundary at lower end points after voiced than after voiceless stops, whereas they had the reverse order of boundaries for the mid-low distinction (fig. 2). Taking the considerations above into account, the results can be interpreted to reflect the operation of a compensatory effect for consonant voicing, if it may be assumed that the listeners used different aspects of the Fo variation as cues to their identification of tones, namely the end point value, or general Fo level, for the high-mid tone distinction and the Fo slope for the mid-low tone distinction. The explanation ventured here is, of course, highly speculative, and calls for further substantiation, but one point speaks in favor of perceptual compensation - or, rather, against non-compensation -, namely the fact that the tonal boundaries were affected by consonant voicing at all. If no compensation had taken place, the boundaries should have been expected to remain constant irrespective of the voicing state of the preceding stop consonant.

In the preceding paragraphs experimental data have been reviewed relating to the questions whether or not segmentally determined Fo perturbations are perceptually compensated for, and whether the two types of perturbation under consideration here receive the same or different treatments in the speech perception process. If the explanation of the results of Abramson and Erickson (1978) attempted above is tenable, two points stand out being of major interest in relation to these questions. One is that in the perception of prosodic categories (stress position and tone), the Fo perturbation due to vowel height and the perturbation due to obstruent voicing both seem to be compensated for. This is in agreement with the view expressed above that the two types of perturbation should not necessarily be expected to be treated differently in speech perception. The other point is that the consonantal Fo perturbation is evidently used by the listeners as a cue to consonant identity, i.e. it is 'heard' by the listeners, whereas, when the identification of tonal categories is concerned, it appears to be compensated for perceptually.

Before any conclusions are drawn as to the implications of the data reviewed above it is necessary to consider one issue which - as mentioned earlier - has not been investigated experimentally, namely the interaction between intrinsic Fo differences among vowels and the perception of vowel height. The experiment to be reported in the following was intended to examine this issue, that is, more specifically, to give an answer to the question: can Fo act as a cue to the perception of vowel height in the same sense as it does to the perception of consonant voicing?

V. Fo AS A CUE TO VOWEL HEIGHT, A PILOT EXPERIMENT

In order to investigate this question a series of words of the type [ˈb̥V:ðə] was synthesized on an OVE III speech synthesizer. The stressed vowel was varied in seven equal logarithmic steps from /u:/ to /o:/. The formant frequencies at the extremes of the vowel quality continuum were: F_1=225 Hz, F_2=825 Hz, F_3= 2100 Hz for /u:/, and F_1=325 Hz, F_2=650 Hz, F_3= 2400 Hz for /o:/. Each vowel quality was combined with three essentially parallel Fo contours at different levels as shown in figure 3. The differences between adjacent Fo levels were 6 Hz, which approximately corresponds in magnitude to the differences in intrinsic Fo between /u:/ and /o:/ reported for Danish by Reinholt Petersen (1978) (see above). The 21 stimuli were arranged in 10 different random orders in an identification test in which the listeners were to decide whether they heard an /u:/- or an /o:/-word. The test was taken by a group of 7 Danish listeners, all staff members or students at the Institute of Phonetics.

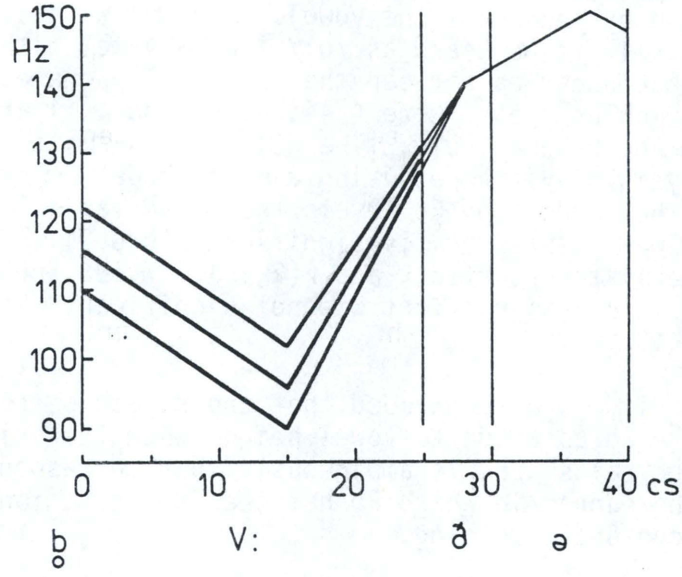

Figure 3

Fo contours of the synthetic ['b̥V:ðə] stimuli used for the identification test.

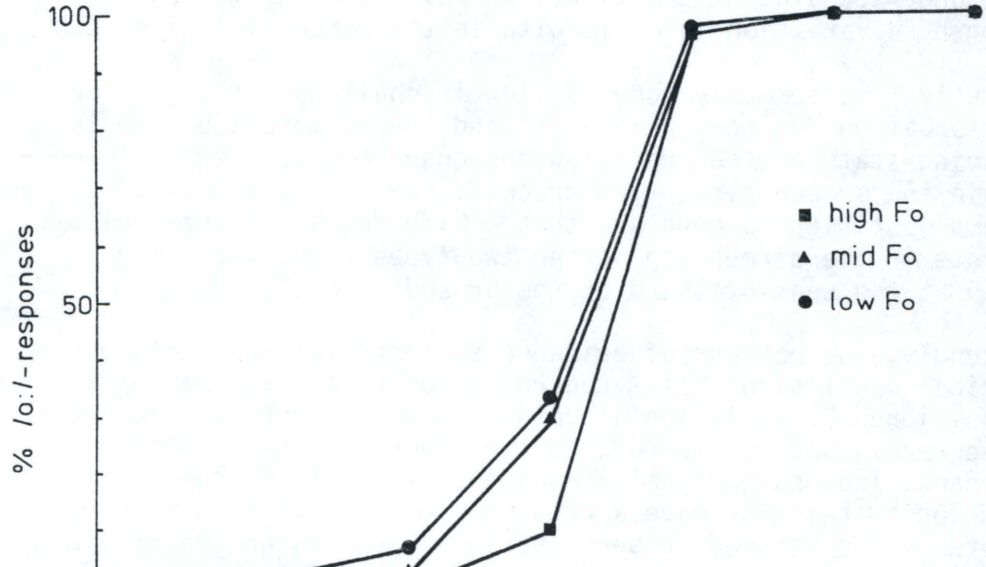

Figure 4

Identification functions (percentage of /o:/-responses) for all subjects pooled under high, mid, and low Fo conditions.

The identification functions for all subjects pooled are displayed in figure 4. It is seen that the response distributions are influenced by the Fo of the vowel. A vowel with a lower Fo is more likely to be heard as /o:/ than a vowel with higher Fo. The mean boundaries between the two vowel phonemes, expressed in stimulus steps, were 4.44, 4.21, and 4.11 at high, mid, and low Fo, respectively. The data were submitted to a two-way analysis of variance, using a mixed model, where listeners were the random and Fo levels the fixed variable. Both main effects were statistically significant ($p<0.01$), $F(6,189) = 4.5$ for the listener effect, and $F(2,12) = 7.536$ for the Fo effect. The interaction effect was not significant $F(12,189) = 0.867$, $p>0.10$.

On this basis it can be concluded that the speech perception system can use Fo as a cue to vowel height when the primary cue, i.e. spectral shape, is ambiguous. This corresponds closely to the manner in which Fo has been shown to function as a cue to consonant voicing.

VI. CONCLUSIONS

The results of the experiment reported above on Fo as a cue to vowel height give additional support to the interpretations outlined previously on the basis of the research surveyed in the present paper. Thus, a pattern of evidence seems to emerge of which two points are of particular significance to the understanding of the manner in which segmentally conditioned Fo variation is dealt with in the perception of speech.

Firstly, the commonly adopted view is challenged that the Fo perturbation due to vowel height and the perturbation due to voicing state of the preceding consonant are treated differently in the speech perception process. On the contrary, it seems justified to conclude that a high degree of parallelism exists in the perception of the two types of Fo variation, both at the segmental and at the prosodic level.

Secondly, the pattern of evidence suggests the seemingly paradoxical conclusion that - regardless of type - the segmentally conditioned Fo variation is both compensated for, at the prosodic level, and is 'heard', at the segmental level. This is, perhaps, less paradoxical if we consider that the speech perception system must have a knowledge of the existence of segmentally conditioned Fo perturbations (or the constraints in the speech production apparatus responsible for them), otherwise it would not be able either to use them as cues for segments or to compensate for them. Under these circumstances it seems obvious to hypothesize that the perceptual system may be sufficiently flexible to apply this knowledge at either level, the segmental or the prosodic, depending on the actual demands on the system. That is, it focuses selectively upon the level where the acoustic input is ambiguous in relation to the identity of the linguistic categories and/or where the greater importance is attached to the correct categorization.

Such a task-dependent behavior of the perceptual system can be considered likely to come out most clearly under the experimental conditions used in the studies considered, which are typically designed to examine one level at a time. In natural speech, cases may well be imagined where the acoustic cues to category membership are ambiguous at the segmental and at the prosodic level at the same time, resulting in misperceptions at one of the levels, or both. However, the probability of occurrence of such cases will, perhaps, be rather small, and in natural speech - as opposed to experimental conditions - the speech perception process will have recourse to the linguistic context in a wide sense.

Finally, it should be emphasized that the conclusions arrived at here clearly call for further experimental substantiation. Indeed, except for the research into the function of Fo as a cue to obstruent voicing, remarkably little experimental work has been done relating to the role of segmentally conditioned Fo variation in the perception of speech. The present paper may, hopefully, provide a framework for such future experimentation.

REFERENCES

Abramson, A.S. 1975: "Pitch in the perception of voicing states in Thai: Diachronic implications", *Status Rep. Speech Res., Haskins Labs. 41*, p. 165-174

Abramson, A.S. and Erickson, D.M. 1978: "Diachronic tone splits and voicing shifts in Thai: Some perceptual data", *Status Rep. Speech Res., Haskins Labs. 53*, p. 85-96

Abramson, A.S. and Lisker, L. 1985: "Relative power of cues: Fo shift versus voice timing", In: *Linguistic Phonetics* (ed.: Fromkin, V.), p. 25-33

Bernstein, L.E. 1983: "Perceptual development for labeling words varying in voice onset time and fundamental frequency", *J. Phonetics 11*, p. 383-393

Chuang, C.-K. and Wang, W.S.-Y. 1978: "Psychophysical pitch biases related to vowel quality, intensity difference, and sequential order", *J. Acoust. Soc. Am. 64*, p. 1004-1014

Di Cristo, A. 1985: *De la microprosodie à l'intonosyntaxe*, (Aix-en-Provence)

Flanagan, J.L. and Saslow, M.G. 1958: "Pitch discrimination of synthetic vowels", *J. Acoust. Soc. Am. 30*, p. 435-442

Fujimura, O. 1971: "Remarks on stop consonants - synthesis experiments and acoustic cues", In: *Form and Substance* (eds.: Hammerich, L.L., Jakobson, R., Zwirner, E.), p. 221-232

Haggard, M., Ambler, S., and Callow, M. 1970: "Pitch as a voicing cue", *J. Acoust. Soc. Am. 47*, p. 613-617

Haggard, M., Summerfield, Q., and Roberts, M. 1981: "Psychoacoustical and cultural determinants of phoneme boundaries: Evidence from trading Fo cues in the voiced-voiceless distinction", *J. Phonetics 9*, p. 49-62

Hombert, J.-M. 1976a: "Phonetic explanations of the development of tones from prevocalic consonants", *Working Papers in Phonetics, UCLA 33*, p. 23-39

Hombert, J.-M. 1976b: "Development of tones from vowel height", *Working Papers in Phonetics, UCLA 33*, p. 55-66

Hombert, J.-M. 1978: "Consonant types, vowel quality, and tone", In: *Tone: A Linguistic Survey* (ed.: Fromkin, V.), p. 77-111

Hombert, J.-M., Ohala, J.J., and Ewan, W.G. 1976: "Tonogenesis: Theories and queries", *Report of the Phonology Laboratory, Berkeley 1*, p. 47-77

Klatt, D.H. 1973: "Discrimination of fundamental frequency contours in synthetic speech: Implications for models of pitch perception", *J. Acoust. Soc. Am. 53*, p. 8-16

Lea, W.A. 1973: "Segmental and suprasegmental influences on fundamental frequency contours", *Consonant Types and Tone, Southern California Occasional Papers in Linguistics 1* (ed.: Hyman, L.M.), p. 17-69

Lindblom, B. and Studdert-Kennedy, M. 1967: "On the role of formant transitions in vowel recognition", *J. Acoust. Soc. Am. 42*, p. 830-843

Massaro, D.M. and Cohen, M.M. 1976: "The contribution of fundamental frequency and voice onset time to the /zi/-/si/ distinction", *J. Acoust. Soc. Am. 60*, p. 704-717

Massaro, D.M. and Cohen, M.M. 1977: "Voice onset time and fundamental frequency as cues to the /zi/-/si/ distinction", *Perc. Psychophys. 22*, p. 373-382

Mohr, B. 1971: "Intrinsic variations in the speech signal", *Phonetica 23*, p. 65-93

Ohala, J.J. 1973: "Explanations for the intrinsic pitch of vowels", *Monthly Internal Memorandum, Phonology Laboratory, University of California, Berkeley*, p. 9-26

Ohala, J.J. 1978: "The production of tone", In: *Tone: A Linguistic Survey* (ed.: Fromkin, V.), p. 5-39

Ohala, J.J. 1982: "Physiological mechanisms underlying tone and intonation", *The XIII International Congress of Lin-*

guists, Working Group on Intonation, preprints, p. 1-12

Ohde, R.N. 1984: "Fundamental frequency as an acoustic correlate of stop consonant voicing", *J. Acoust. Soc. Am. 75*, p. 224-230

Ohde, R.N. 1985: "Fundamental frequency correlates of stop consonant voicing and vowel quality in the speech of preadolescent children", *J. Acoust. Soc. Am. 78*, p. 1554-1561

Reinholt Petersen, N. 1978: "Intrinsic fundamental frequency of Danish vowels", *J. Phonetics 6*, p. 177-189

Reinholt Petersen, N. 1983: "The effect of consonant type on fundamental frequency and larynx height in Danish", *Ann. Rep. Inst. Phon., Univ. Cph. 17*, p. 55-86

Rosenvold, E. 1981: "The role of intrinsic Fo and duration in the perception of stress", *Ann. Rep. Inst. Phon., Univ. Cph. 15*, p. 147-166

Silverman, K. 1984a: "What causes vowels to have intrinsic fundamental frequency?", *Cambridge Papers in Phonetics and Experimental Linguistics 3*, p. 1-15

Silverman, K. 1984b: "Fo perturbations as a function of voicing of prevocalic and postvocalic stops and fricatives, and of syllable stress", *Proceedings of the Institute of Acoustics, (Autumn conference, Windermere) 6*, p. 445-452

Stoll, G. 1984: "Pitch of vowels: experimental and theoretical investigation of its dependence on vowel quality", *Speech Communication 3*, p. 137-150

Terhardt, E. 1979: "Calculating virtual pitch", *Hearing Research 1*, p. 155-182

Terhardt, E., Stoll, G., and Seewann, M. 1982a: "Algorithm for extraction of pitch and pitch salience from complex tonal signals", *J. Acoust. Soc. Am. 71*, p. 679-688

Terhardt, E., Stoll, G., and Seewann, M. 1982b: "Pitch of complex signals according to virtual pitch theory: tests, examples, and predictions", *J. Acoust. Soc. Am. 71*, p. 671-678

Thorsen, N. 1984: "Variability and invariance in Danish stress group patterns", *Phonetica 41*, p. 88-102.