Journal of Phonetics (1995) **23**, 349–366

966

The universality of intrinsic F_0 of vowels

D. H. Whalen

Haskins Laboratories, 270 Crown St., New Haven, CT 06511, U.S.A.

Andrea G. Levitt

Haskins Laboratories, 270 Crown St., New Haven, CT 06511, and French Department, Wellesley College, Wellesley, MA 02181, U.S.A.

Received 18th April 1994, and in revised form 24th January 1995

The tendency for high vowels such as [i] and [u] to have higher fundamental frequencies (F_0s) than low vowels such as [a] has been found in every language so far in which it has been sought. This includes 31 languages representing 11 of the world's 29 major language families (as defined by Crystal, 1987). While the size of the intrinsic F_0 (IF_0) effect varies from study to study, the differences seem to derive from differences in the study design, especially in the number of subjects. The effect appears larger for female speakers when expressed in Hz, but it is, instead, larger for males when the results are expressed in semitones. The size of the language's vowel inventory did not significantly affect the size of IF_0. One other universal, though, is that the effect disappears at the low end of a speaker's F_0 range. The consistency of intrinsic F_0 across languages argues that the effect is truly intrinsic; that is, it is not a deliberate enhancement of the signal but rather a consequence of successfully forming a vowel. © 1995 Academic Press Limited.

1. Introduction

Although all spoken human languages make use of sounds produced by the vocal tract, there is a great deal of latitude in the sounds selected. Every language uses vowel sounds, for example, but the number of distinctive vowels ranges from a low of 2 for Margi or Ubykh (Ladefoged & Maddieson, 1990) to a high of 24 (for !Xũ) in the UCLA database of 317 languages (Maddieson, 1984). For languages with the same number of vowels, there is a great range of combinations of vowel qualities selected. These vowels are also the primary locus for lexical suprasegmental distinctions which exist in many languages, such as stress, pitch accent, and tones. Such differences are among the features that make languages distinct.

One phonetic feature that has been found to accompany vowels is "intrinsic F_0" or "intrinsic pitch" (IF_0, from here on). This is the tendency for the high vowels, such as [i] and [u], to have a higher fundamental frequency than the low vowels, such as [a] and [æ]. IF_0 was first noticed for German (Meyer, 1896–7) and has since been found in every language that has been examined for it. Our goal in the present paper is to examine the size of the effect in all the languages reported so far to see whether there is any difference that does not seem to be due to the factors of

0095-4470/95/030349 + 18 $12.00/0

experimental procedure. The features we will study will include not just language but also speaker sex and the size of the language's vowel inventory. The survey will also provide an estimate of the range of variability that can be expected in measuring IF_0.

The mechanism behind this effect is the object of considerable debate. We will not give a full history of the different explanations, since such surveys exist elsewhere (DiCristo, Hirst, & Nishinuma, 1979; Shadle, 1985; Silverman, 1987; Sapir, 1989; Fischer-Jørgensen, 1990). All of the explanations surveyed there have assumed that IF_0 is an automatic consequence of articulation, most likely the pull of the tongue on the laryngeal system, or acoustics. (Steele (1986) argues that there must be a contribution of subglottal pressure.) However, more recently there have been proposals that IF_0 is simply another deliberate change in F_0 that is introduced in the signal to enhance the differences between vowel categories (Diehl & Kluender, 1989a,b; Diehl, 1991; Kingston, 1993).

The consistency of the effect across the world's languages can provide us with some indication of whether an automatic or deliberate process is more likely. If we find differences among the different languages, it would be likely that the degree of IF_0 is another variable that languages choose, just as they choose their inventory or their suprasegmental category. If, on the other hand, there seems to be little change in the size of IF_0, then we would expect that, whatever mechanism is responsible, it is truly intrinsic, and occurs as part of the production of vowels in any language. Such a conclusion can only be provisional, of course, since it is (in the statistical sense) the null hypothesis. Nonetheless, a failure to observe a difference in the magnitude of the IF_0 effect would pose a challenge for any enhancement explanation of the phenomenon.

2. The published results

Our data come from various published sources, listed along with the IF_0 values in Table I. These include all of the articles published in journals, proceedings and working papers that we were able to locate. In many cases, IF_0 is not directly assessed in the work cited, but the numbers from which IF_0 could be calculated were given. Table II lists information about the languages, including the size of the vowel inventory.[1]

Some of the descriptions are incomplete, and so question marks appear in some places. In some cases, it was difficult to find the appropriate measures. For example, the Serbo-Croatian results (Ivic & Lehiste, 1965) have been cited a number of times as showing IF_0, and yet there is no condition in which each of the vowels appears with the same pitch accent. We have tried to minimize the effect of accent for this study of Serbo-Croatian by averaging the values at the beginning of the vowel, a point at which the F_0 excursion for the pitch accent should not be as advanced as later. Only four words were measured (brȁtu, grâdu, Mári, sȅlu), which is a smaller set than would be best. Any perturbation due to consonant voicing should be

[1] After this paper was submitted, we became aware of Zhu's (1994) report on Shanghai. He reports values for six men and five women as follows (taken from the end of tone 2, as we did for King *et al.*, 1987): /u/: 131/244; /i/: 130/231; /a/: 118/217.

Zhu suggests that IF_0 occurs even with low tones, as long as we look exclusively at the "target" of the tone. Readers are referred to his work for a more detailed explication.

TABLE I. Values for intrinsic F_0 from published sources. All F_0 values are in Hz. Values in italics were not used in the statistical test. Those studies that lacked values for [u] were also not used for the statistical test, but are included in Fig. 1. If results were combined for males and females, the sex is listed as "B". In the "Context", "V" stands for vowel and "S" for sentence

Language, source	#, Sex of S's	Context	F_0 for [u]	F_0 for [i]	F_0 for [a]
English					
Crandall, 1925	4 M	Isolated Vs	140	136	133
	4 F	Isolated Vs	270	252	234
Taylor, 1933	8 M	Isolated real words	152	149	132
	9 F	Isolated real words	323	320	298
Black, 1949	16 M	Isolated CVC	153	146	133
Peterson & Barney, 1952	33 M	Isolated hVd	141	136	124
	28 F	Isolated hVd	231	235	212
House & Fairbanks, 1953	10?	?	130	128	118
Lehiste & Peterson, 1961	5 M	?	124	129	120
Peterson, 1961	4 M	Isolated V	128	124	119
	3 F	Isolated V	253	250	212
Atkinson, 1973	5 M	Sentences	128	132	114
Fox, 1982	8 F	Isolated hVd	240	242	234
	8 M	Isolated hVd	145	148	140
Shadle, 1985	2 F	Sentence, first 3 positions	225	225	215
		Sentence, last position	*175*	*174*	*175*
	2 M	Sentence, first 3 positions	116	115	108
		Sentence, last position	*93*	*93*	*90*
Zawadzski & Gilbert, 1989	5 M	?	208	206	188
Nittrouer, McGowan,	4 M	Nonsense CV in S	139	138	132
Milenkovic & Beehler, 1990	4 F	Nonsense CV in S	215	210	197
Higgins *et al.*, 1994	11 M	Isolated [pa] or [pi]	—	125	120
		Sentences w/[pap] or [pip]	—	*126*	*121*
	10 F	Isolated [pa] or [pi]	—	206	196
		Sentences w/[pap] or [pip]	—	*212*	*194*
Hillenbrand *et al.*, 1995	45 M	Isolated hVd	143	138	123
	48 F	Isolated hVd	235	227	215
Dutch					
Koopmans-van Beinum, 1980	2 M	Monosyllabic words	153	149	126
	2 F	Monosyllabic words	217	229	205
van Son, 1993	1 M	Running text (+accent only)	197	193	186
German					
Meyer, 1896–7	1 M	Isolated Vs	121	108	87
Mohr, 1971	1 M	Initial V	122	121	116
Neweklowsky, 1975	2 M	Isolated pVp (short V)	139	137	124
(all are Viennese)	1 F	Isolated pVp (short V)	219	222	210
Antoniadis & Strube, 1981	3 M	CVCə (long V)	132	131	125
		CVCə (short V)	*135*	*133*	*129*
Traunmüller, 1982	12 M	Nonsense CV in S	126	133	105
Möbius, Zimmermann & Hess, 1987	1 M	Real words in carrier (long V)	116	123	107
Iivonen, 1989	5 M	Isolated words (long V)	95	90	82
Iivonen, 1989 (Viennese)	5 M	Isolated words (long V)	140	138	106
Fischer-Jørgensen, 1990	5 M	dVdə in carrier S (long V)	—	125	112
	1 F	dVdə in carrier S (long V)	—	184	167

TABLE I. (*Continued*)

Language, source	#, Sex of S's	Context	F_0 for [u]	F_0 for [i]	F_0 for [a]
Danish					
Reinholt Petersen, 1978[a]	2 F	CVCV:CV in S (middle syll)	220	212	182
	3 M	CVCV:CV in S (middle syll)	122	120	104
Swedish					
Fant, 1959	7 M	Isolated Vs	127	128	124
	7 F	Isolated Vs	222	218	215
French					
Rossi & Autesserre, 1981	4 M	Isolated Vs	142	140	131
DiCristo, 1982	1 F	?	243	235	226
	3 M	?	134	132	124
Lavoie, 1994 (Quebecois)	2 M	Short phrases	198	199	185
	2 F	Short phrases	300	289	266
Italian					
Pettorino, 1987	1 M?	Isolated real words	158	159	143
Greek					
Samaras, 1972	2 M	CVC	148	149	137
Russian					
Mohr, 1971	1 M	Initial V	127	126	121
Bolla, 1981	1 M	Isolated real words	135	136	116
Polish					
Steffen-Batóg, 1970	6 M	Isolated VCV	153	153	150
Serbo-Croatian					
Ivic & Lehiste, 1965	12 B	Words in initial and medial position in S, beg. of syll	183	217	194
		Final S position, beg. of syll	*129*	*138*	*159*
Standard Croatian					
Bakran & Stamenkovic, 1990	3 M	?	108	110	98
	3 F	?	177	197	165
Lithuanian					
Pakerys, 1982	7 M?	Isolated words—stressed	149	152	139
		—unstressed	*128*	*126*	*125*
Hindi					
Schiefer, 1987	1 F	Isolated real words	252	238	205
Gujarati					
Dave, 1967	3 M	Isolated Vs, words, phrases	136	145	128
Finnish					
Vilkman, *et al.*, 1989	1 M	Isolated Vs	119	118	114
Hungarian					
Tamas, 1976	? M	stressed vowels	137	144	120
		unstressed vowels	*103*	*90*	*85*
	? F	stressed vowels	231	235	205
		unstressed vowels	*195*	*185*	*170*
Korean					
Han & Weitzman, 1967	1 M	Nonsense CV	185	183	166
	1 F	Nonsense CV, CVC	329	316	317
Kim, 1968	? F	?	279	277	268
Japanese					
Homma, 1973	1 F	Words?	342	350	328
Nishinuma, 1979	5 M	2nd syll of 4 syll words	148	142	136

TABLE I. (*Continued*)

Language, source	#, Sex of S's	Context	F_0 for [u]	F_0 for [i]	F_0 for [a]
"Chinese"					
Mohr, 1971	1 M	Initial V	148	150	147
Mandarin					
Shi & Zhang, 1987	5 M	Tone 1 (=high F_0)	181	175	154
		Beg. of Tone 2 (=low F_0)	*117*	*118*	*111*
		End of Tone 2 (=high F_0)	*168*	*167*	*151*
		Beg. of Tone 3 (=mid F_0)	*112*	*113*	*108*
		End of Tone 3 (=low F_0)	*90*	*89*	*83*
		Beg. of Tone 4 (=high F_0)	*206*	*197*	*175*
		End of Tone 4 (=low F_0)	*105*	*97*	*97*
Shi & Zhang, 1987	5 F	Tone 1 (=high F_0)	307	297	276
		Beg. of tone 2 (=low F_0)	*209*	*205*	*198*
		End of Tone 2 (=high F_0)	*289*	*265*	*255*
		Beg. of Tone 3 (=mid F_0)	*218*	*219*	*227*
		End of Tone 3 (=low F_0)	*172*	*169*	*171*
		Beg. of Tone 4 (=high F_0)	*335*	*312*	*302*
		End of Tone 4 (=low F_0)	*184*	*180*	*187*
Taiwanese					
Zee, 1980	3 M	Sentences, High Tone	157	157	143
		Sentences, Low Tone	*100*	*99*	*102*
Shanghai					
King, Ramming, Schiefer & Tillmann, 1987	1 M	Isolated words, Tone 2, end	156	163	142
Kammu					
Svantesson, 1988	1 M	Words in S: High tone	150	148	129
		: Low tone	*115*	*116*	*104*
Paraok					
Svantesson, 1993	1 M	Isolated words	159	156	142
	1 F	Isolated words	293	305	250
Vietnamese					
Han, 1969	1 F	Vs in carrier S: High tone	308	309	288
		End of falling tone	*185*	*189*	*194*
	1 M	Vs in carrier S: High tone	143	141	136
		End of falling tone	*116*	*118*	*107*
Thai					
Gandour & Maddieson, 1976	1 M	Isolated nonsense CV, middle of high tone	148	147	138
Malagasy					
Rakotofiringa, 1968	1? M?	?	124	123	122
Rakotofiringa, 1982	18 M	Isolated words	122	119	116
	11 F	Isolated words	238	233	225
Yoruba					
Hombert, 1977	2?	Isolated V, high tone	177	179	170
		Isolated V, mid tone	*152*	*153*	*147*
		Isolated V, low tone	*121*	*119*	*120*
Itsekiri					
Ladefoged, 1968	1 M	?	(reports a 5 Hz difference)		
Hausa					
Pilszczikowa-Chodak, 1972	1 M	Isolated words	122	141	117

TABLE I. (*Continued*)

Language, source	#, Sex of S's	Context	F_0 for [u]	F_0 for [i]	F_0 for [a]
Navaho					
de Jong & McDonough, 1993	4 F	Words, High tone	—	209	188
		Words, Low tone	—	*181*	*180*
	2 M	Words, High tone	—	133	138
		Words, Low tone	—	*129*	*129*

TABLE II. Information about the languages. Family affiliation is taken from Crystal (1987). The number of vowels represents the number of distinctive vowel qualities in the articulatory space (tongue height, front/back, rounding). Thus vowels that differed distinctively in duration but had the same quality were considered to represent one vowel, not two. Distinctive nasality was also ignored. See text for the rationale for this decision

Language	Family	Sub-family	# of Vs
English	Indo-European	W. Germanic	12
Dutch	Indo-European	W. Germanic	12
German	Indo-European	Germanic	14
Danish	Indo-European	N. Germanic	12
Swedish	Indo-European	N. Germanic	18
French	Indo-European	Italic	12
Italian	Indo-European	Italic	7
Greek	Indo-European	Greek	5
Russian	Indo-European	Slavic	5
Polish	Indo-European	Slavic	6
Serbo-Croatian	Indo-European	Slavic	5
Standard Croatian	Indo-European	Slavic	5
Lithuanian	Indo-European	Balto-Slavic	11
Hindi	Indo-European	Indo-Iranian	14
Gujarati	Indo-European	Indo-Iranian	8
Finnish	Uralic	Finno-Ugric	16
Hungarian	Uralic	Finno-Ugric	10
Korean	Korean		18
Japanese	Japanese		5
"Chinese"	Sino-Tibetan	Sinitic	5?
Mandarin	Sino-Tibetan	Sinitic	5?
Taiwanese	Sino-Tibetan	Sinitic	6
Shanghai	Sino-Tibetan	Sinitic	12
Kammu	Austro-Asiatic	Mon-Khmer	10
Paraok	Austro-Asiatic	Mon-Khmer	9
Vietnamese	Austro-Asiatic	Mon-Khmer	11
Thai	Tai	S.-Western	9
Malagasy	Austronesian	Western	4
Yoruba	Niger-Congo	Kwa	7
Itsekiri	Niger-Congo	Kwa	7
Hausa	Afro-Asiatic	Chadic	5
Navaho	Na-Dené	Athabaskan	4

minimal, since voicing is different for only one of the syllables. These Serbo-Croatian values would be more comparable to the others if we had instances of the different vowels with the same accent. Because of these difficulties, these numbers are not included in the statistical tests reported below. The numbers for Navaho (de Jong & McDonough, 1993) are not represented directly in that paper and have been supplied by the authors. Similarly, the F_0 values for Quebecois French were not in the published version and were supplied by the author (Lavoie, 1994). The "Chinese" of Mohr (1971) is presumably Mandarin (and is so treated in the analysis of variance), but has been listed as "Chinese" in the table.

We restricted our analysis to the vowels [i], [u], and [a] (or [ɑ]). These include the two dimensions that have been examined most, vowel height and front/back. They are present in about 80% of the world's languages (Maddieson, 1984), and occur in almost all the languages examined: only Navaho lacked one of the vowels, namely [u]. However, the high back vowel in the Tokyo dialect of Japanese (the one studied in both references here) is an [ɯ] rather than [u]. There were two other studies that lacked [u] values, simply because they were not measured. One was for English (Higgins, Netsell, & Schulte, 1994) and one for German (Fischer-Jørgensen, 1990). Our focus on these three vowels is not intended to deny that there is gradation between high and low or that the other dimensions of rounding, tense/lax, nasalization, advanced tongue root, etc., might not play a part in IF_0. The selection was made solely to equate the language samples as far as was possible.

The numbers come from published studies, and usually there is no individual data given, so it is not possible to perform ordinary statistics on these numbers. We performed an ANOVA on the averaged results for most of the numbers in Table I. Italicized numbers were not used, since they represent different phonetic environments. They are included as comparisons to the main environment studied in the statistical analysis. We also excluded from the statistical analysis the studies which did not have measurements for [u]. Each study was weighted by the number of subjects measured, so that they contributed to the means in proportion to the amount of data represented. We performed separate tests for a number of factors: vowel ([i] versus [u] versus [a]), front/back ([i] versus [u]), sex, language, English versus the other languages, and languages compared by vowel inventory size. Separate analyses were conducted because there were not enough studies to have anything like equal cell sizes if we included more than one factor. We also performed each test twice, once on the Hz values and once on the values expressed in semitones. The use of semitones helps equate the differences found for the males and females. Since the semitone scale relates one frequency to another, we needed to have a reference frequency. We chose 1 Hz, so that the denominator would fall out, and the simple formula:

$$\text{semitone} = 1/\log(2) \times 12 \times \log(\text{Hz})$$

could be used. The values that we were able to use are somewhat less than optimal, since they are semitone transforms from the mean expressed in Hz, rather than a mean of the values expressed as semitones. But since the majority of studies did not give individual values, this was our only choice.

Our first analysis looked at a single factor with three levels, namely, a vowel effect comparing [u]/[i]/[a]. The vowel effect was highly significant

($F(2, 150) = 246.67$; $p < 0.0001$). Overall, the means were 177.4 Hz for [u], 174.9 Hz for [i], and 160.9 Hz for [a]. This translates into a 15.3 Hz difference for IF_0 across all the languages. The semitone analysis was also highly significant ($F(2, 150) = 231.51$; $p < 0.0001$). Those means were 88.90, 88.66, and 87.13 semitones for [u], [i] and [a] respectively. That difference is 1.65 semitones. As expected, then, IF_0 is a highly significant effect.

Front/back (comparing just [i] and [u]) was not a significant factor, either in the Hz analysis ($F(1, 76) = 2.48$; ns) or the semitone analysis ($F(1, 76) = 3.16$, $p < 0.10$). As can be seen from the means in the previous paragraph, there is a 2.5 Hz (0.24 semitone) difference between the front and back vowels across all the languages. Inspection of Table I will show that some languages seem to have large differences between the front and back vowels. For example, English [u] is reliably higher in F_0 than [i] as measured across the studies. The 184.7 Hz of [u] is reliably (3.6 Hz) higher than the 181.1 of [i] ($F(1, 20) = 12.73$; $p < 0.01$) as is the 0.36 semitone difference in the semitone analysis ($F(1, 20) = 15.19$; $p < 0.001$). While this may indicate that there is a genuine difference in the front and back vowels for English, this analysis would be more reliable if done with the individual measurements rather than the means across experiments. The amount of variability that is present within experiments is not carried across when only the means are used. A comparison case is German, for which we have the next largest data set. In Table I, six of the nine entries for German also show a higher value for [u] than [i], but one of the studies that has higher values for [i] than [u] also had a large number of subjects (Traunmüller, 1982), and so the overall effect is not significant and, in fact, shows [i] being 0.6 Hz higher than [u]. Again, it does not seem impossible that a language could adopt an articulation for one vowel or another that would modify its IF_0, but there is no indication of a universal difference between [i] and [u]. It is possible that the English differences are real, and therefore that some languages have effects that other languages lack, indeed, that other languages have in the other direction. The effect in English is one fourth the size of the height effect and will consequently be harder to verify. It is also likely that some of the apparent differences are simply the result of the inevitable sampling error.

For the remainder of the analyses, we will collapse across [i] and [u] and subtract the value for [a], giving us a simple test for the significance of the difference in height. The test for the means (i.e., whether there was an IF_0 effect) was always significant, and so we will not report the numbers. In this way, the factors that we examine from here on out will directly indicate whether there was an effect of that factor on the size of the IF_0 difference.

For the analysis by sex, we excluded the three studies that did not separate out the two sexes (House & Fairbanks, 1953; Ivic & Lehiste, 1965; Hombert, 1977). The overall IF_0 effects were 13.9 Hz for the males and 15.4 for the females. This difference in magnitude was significant ($F(1, 72) = 5.87$; $p < 0.05$). It would appear from these numbers that women have a larger effect than men. But they also have a higher overall F_0, so the greater magnitude of their effect represents, in fact, a smaller percent increase. This fact is accounted for in the semitone analysis. That analysis, though, also shows an effect ($F(1, 72) = 6.53$; $p < 0.05$), but now the difference is larger for the males than the females (1.84 semitones for the males and 1.34 for the females). It appears that there is a difference in the size of IF_0 by the sex of the speaker. As with the front/back difference, though, this indication must be treated with caution since it was, of necessity, based on the means across experiments and not, as would be optimal, on the individual results. Thus, the lack of a contribution of individual variability may have made a small, random difference appear significant. This aspect of the data will require further elaboration before we can decide whether there is anything that needs explaining. However, the study with the largest number of subjects (Hillenbrand, Getty, Clark, & Wheeler, 1995) had a sizable difference between males and females, so it may simply be a factor that requires substantial evidence before it is apparent. While it is conceivable that the lowering of the male larynx during puberty could be a factor in such a difference, no explanation is immediately obvious.

The analysis by language did not reveal an IF_0 difference for any one language that was statistically different from the others ($F(27, 48) = 1.27$; ns for the Hz analysis; $F(27, 48) = 1.61$; ns for the semitone analysis). Since approximately half of our results came from English, it was of interest to see whether the English results patterned with the others. We did this by classifying each IF_0 measurement as being either English or non-English. This also showed no difference ($F(1, 74) = 2.39$; ns for the Hz analysis; $F(1, 74) < 1$; ns for the semitone analysis). Thus, there is no evidence that English stands out from the others in terms of the size of the IF_0 effect. This test, besides its weakness as accepting the null hypothesis, must also be treated with caution because so many of the languages were poorly represented. If there were just one language that showed an effect of a different size, it would be very hard to detect without measuring many more subjects. The one language (English) that we do have many measurements for, though, seems to be typical.

Finally, we separated the languages by the size of their vowel inventory. Since we have such a large amount of data on English, the 12-vowel systems are automatically overrepresented. We thus classed them by themselves. For the rest, we divided them into small (4–5), medium (6–11) and large (13 or more) vowel inventories. The inventory size is taken as the number of monophthongal distinctive vowels that differ in quality. In many cases, these same vowels could also differ in length or nasalization, but that was not considered to make a different vowel for this analysis. If length makes a distinction without changing quality, then the long and short were counted as one vowel, while long and short vowels that did differ (such as those in German) were counted as two vowels. This form of counting vowels was adopted under the assumption that any enhancement that might be introduced deliberately would be likely to depend on how crowded the vowel space is. Even if such separable dimensions as length and nasalization are distinctive for a language, they should not influence any use of F_0 in perception, since F_0 is posited to affect primarily the height dimension.

The size of the IF_0 effect was 11.6, 12.6, 16.3 and 15.1 Hz for the small, medium, 12-vowel and large inventories, respectively. These did not differ significantly ($F(3, 72) = 1.56$; ns). The size of the IF_0 effect in semitones was 1.17, 1.33, 1.70 and 1.64 for the small, medium, 12-vowel and large inventories, respectively. These also did not differ significantly ($F(3, 72) = 2.47$; ns). Again, we need a more balanced set of results before we can be sure that these statistics do not hide a small effect, but the current indication is that vowel inventory size also does not affect the size of IF_0.

For some of the 10 tone languages in the survey, there is an additional observation that can be made: The low tones fail to show IF_0. While this is true of all but one of the tone languages in Table I, perhaps the most striking case is the Tone 4 data of

Shi & Zhang (1987). There, within one syllable, the IF_0 effect was present at the onset (the high portion of the F_0 contour) and then absent at the offset (the low frequency). Such results match up with the findings of Shadle (1985) and Ladd & Silverman (1984) for spontaneous speech (for English and German, respectively). In fact, Ladd & Silverman claim that "low pitch, rather than low-stress or phrase-final position per se, is the relevant factor" (1984: 36). Of course, this statement is slightly misleading, since what is important is the relative frequency within a speaker's range rather than absolute F_0. Still, the main point is borne out by the tone languages, that in the lower part of the frequency range, IF_0 disappears. The one case where it does not is Kammu. The high tone has an effect of 20 Hz while the low has an effect of 12 Hz. This is a fairly substantial difference, but the 12 Hz on the low tone is certainly within the range of values found for other studies with just one speaker. Although this language may constitute an exception to the disappearance of IF_0 at the lower part of the F_0 range, it is also possible that the "low" tone is simply low in relation to the high tone, and that a mid-range pitch is used for it. In that case, we would expect to find an IF_0 effect. Further study is needed before deciding this issue, but it is interesting to compare the Kammu results with the Yoruba in Table I. Yoruba has three tones, high, mid and low. The IF_0 effect is 8, 6 and 0 Hz for those, respectively. So the existence of an IF_0 effect for a "lower" (i.e., mid) tone is not out of the question.

The range of variability shown by these studies appears rather large, but there is a tendency, noted above, for the larger studies to approach more closely the mean across studies (without being weighted by number of subjects). Such an outcome is to be expected if the measurements come from a single distribution, but it is not a necessary outcome if some languages differ from the rest. The measured IF_0 does differ across studies, but this appears to be due to typical distributional factors. There are three sources of variability in the published results. The first source is the distribution of IF_0 values for different languages. Since speech is a natural system, we have to expect some variation in even its most consistent aspects. If we analyzed a large number of speakers for each language (and languages truly do not differ), the distribution of IF_0 values would be extremely tight. If languages were selecting from a small range of choices (e.g., either the typical IF_0 or none at all), then we might find a bimodal distribution. The second source of a distribution is in the individual IF_0 values. If we looked closely at values for individual speakers within a language, we would expect to see somewhat larger variability than for the languages themselves, since it is easier for an individual to be extreme than for a whole language community. These first two are sources of variability in the real world. A final source of variability is in our measures of the world. These will also be distributed around the true mean, since any measurement is inherently fallible. This measure will also tend to form tighter distributions as the number of subjects increases.

As a way of judging the combined effects of these three distributions, we have plotted in Fig. 1 the difference between the value for [a] and the average for [i] and [u] from Table I. These include all the lines from Table I that were used in the analysis of variance, along with the values from the studies lacking an [u]. The top panel shows the results for the Hz analysis, and the bottom, for the semitone analysis. The number of subjects in each study is shown on the y-axis. We have plotted the data this way rather than doing a more typical histogram because there

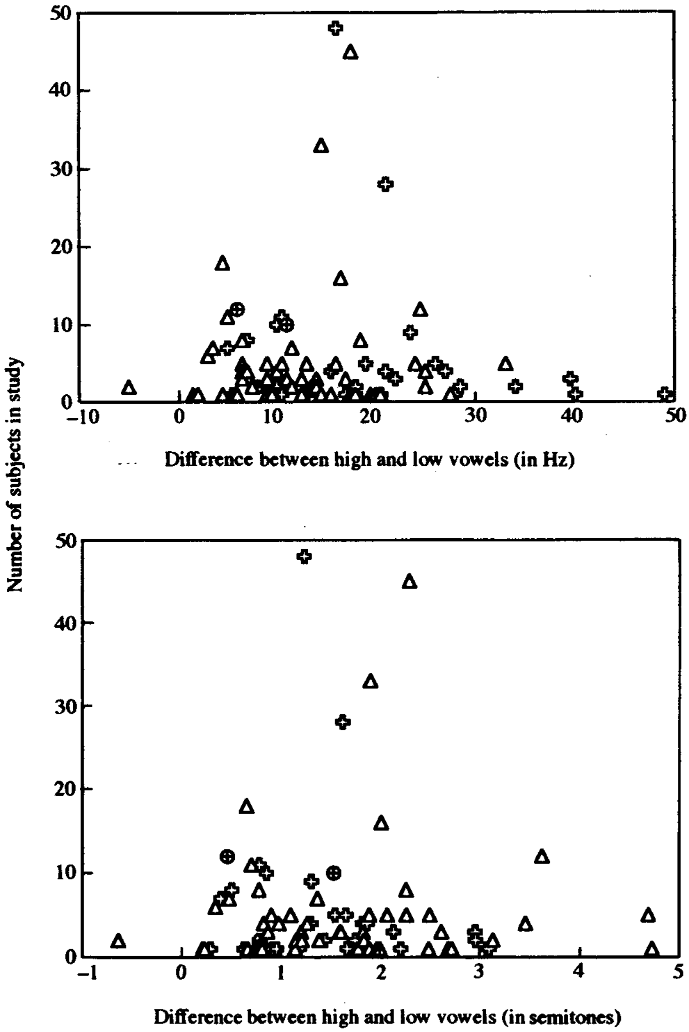

Figure 1. Distribution of the differences between the two high vowels and the low vowel for the studies in Table I. The top panel represents the differences between the values expressed in Hz, and the lower panel, between the values expressed as semitones. Δ, Males; ✚, females; ⊕, both.

are so many gaps in our distribution, in terms of the number of subjects per study. While we certainly do not have a large enough sample to show that the values are distributed normally, there is a tendency for the larger studies to show values close to the overall mean difference of 15.3 Hz. The semitone analysis shows an even tighter distribution, with both ends of the distribution pulling in towards the mean of 1.65 semitones. Again, only a much larger set of results for all of the languages in this set would allow us to test for this tendency statistically, but the distributions shown in Fig. 1 are consistent with an absence of any difference across languages. It is also easy to find where a new set of results would fit in to see whether it conforms to the current pattern.

3. IF_0 as enhancement?

Diehl and colleagues (Diehl & Kluender, 1989a,b; Diehl, 1991) have claimed that IF_0 is an enhancement of the speech signal. Enhancements, in their view, are deliberate attempts by the speaker to make the auditory aspects of speech more salient for the hearer. Perception of vowel height is claimed to be influenced by the difference between F_1 and F_0. IF_0, then, enhances this difference, since a high vowel will have a low F_1. If this low F_1 is accompanied by a high F_0, the difference will be even smaller than before, and thus even more distinct from the large difference for low vowels.

One aspect of other proposed enhancements is that they are optional. For example, the typical vowel systems of the world combine rounding with back vowels, to increase the proximity of F_2 and F_1 (Lindblom, 1986). Yet there are languages, such as the Iroquoian languages, which avoid the use of lip rounding (e.g., Oneida: Lounsbury, 1953: 27). If IF_0 is truly universal and unmodified, it is less likely to be under voluntary control. Of course, even if we had data on every language of the world, it still might be the case that they all happened to choose to enhance the vowels this way, and that the next human language to evolve would dispense with IF_0. In the absence of even one example of a language doing without, however, the enhancement view is tenuous.

The magnitude of the shift in vowel identity that we might expect from the IF_0 differences is rather small. In a survey of such studies, Nearey (1989) found the following size of shifts in the frequency of F_1 for an octave shift in F_0: 16%, 6%, 14%, 21%, 21%, and 26–30% (see his page 2093). As a rough estimate, then, we might expect a 17% shift in effective F_1 for an octave shift in F_0. The IF_0 effect, however, is much smaller than an octave, approximately one seventh that size (around 1.7 semitones), which would lead us to expect about a 2.4% shift in the effective F_1. This value is below the difference limen (of 3–5%) for changes in single formants of vowels (Flanagan, 1955). It is difficult to see why such a complicated change in vowel articulation would be deliberately introduced for such a modest reward, especially for those languages with 4–5 vowels. The enhancement account would seem more likely if there were an effect of vowel inventory size on the size of IF_0, so that more vowels would lead to a greater use of IF_0. The present results showed no such effect, however.

It is also hard to see why tone languages would want to use F_0, which is critical for tone, to enhance vowel identification, especially in a system like Mandarin's in which the vowel system is quite simple (5 vowels). The tones occur with every vowel, and therefore the tonal differences of over 100 Hz appear with each vowel, yet the 5–10 Hz IF_0 effect remains. Perhaps a parsing of the F_0 into its tone and IF_0 components would make this work, but such an effect is indicative of a perceptual system operating on the signal, not the low-level auditory effect that the enhancement theory presupposes. It seems more likely that Mandarin and other tone languages exhibit IF_0 because it is a natural consequence of producing vowels of different heights: Even when speakers change the F_0 of a vowel (within the upper part of the F_0 range, at least), there is a contribution of IF_0 present.

One other aspect of IF_0 that appears to favor a deliberate component is the fact that IF_0 appears in some laryngectomized patients who use a flap over the esophagus to voice their sounds (Gandour & Weinberg, 1980; Pettorino, 1987), but there is a problem with claiming that such evidence means that IF_0 in normal speech is deliberately produced. Even if we assume that these speakers are introducing IF_0 deliberately, we would not know anything more about the normal case: The esophageal speakers could simply be recreating what used to come to them naturally. If they monitor their own productions and find that their [i]'s, for example, are consistently too low in frequency to sound right, then they may deliberately raise the F_0. The IF_0 of esophageal vowels does not indicate that normal IF_0 is deliberate.

If it is important that the vowel judgement be enhanced, it is not clear why speakers would choose not to exhibit IF_0 in the lower portion of their frequency range. While it is certainly the case for English that words that end up in the low range are unstressed and therefore perhaps dispensable, the low tones of tone languages can occur with any word, perhaps the most important one in the sentence. Kingston (1993) has claimed that the lack of an IF_0 effect in the lower frequency ranges constitutes evidence in favor of the enhancement account, but only because the automatic account does not have a ready explanation for it. Conceivably, the arrangement of the larynx makes it very difficult to introduce these differences at low frequencies, and thus the gain from enhancement is judged not to be worth the physiological cost. We will propose below that there is, indeed, something physiologically different about the lower portion of the F_0 range. Certainly, other muscles become involved in the lower frequency range (Hallé, 1994). It would be easier to judge whether this lack of an effect in the low frequencies is also automatic or just a cost decision if we could make quantitative predictions about just how much energy is involved. Unfortunately, to make such estimates would require that we know enough about the laryngeal system that we would know the answer already—we would already know whether IF_0 was a necessary consequence of the way the larynx is employed in speech.

The most direct piece of evidence for the enhancement theory is the fact that cricothyroid (CT) activity increases for high vowels (Autesserre, Roubeau, DiCristo, Chevrie-Muller, Hirst, Lacau, & Maton, 1987; Vilkman, Aaltonen, Raimo, Arajärvi, & Oksanen, 1989; Honda & Fujimura, 1991). Increasing CT activity generally increases F_0. Thus it appears that these subjects were intentionally increasing F_0 for the high vowels. However, the laryngeal system is quite complex, and the increase in CT activity may not be exclusively associated with higher F_0. Vilkman *et al.* (1989), in fact, consider the intentional explanation and reject it in favor of one in which the CT activity is increased to "avoid opening of the cricothyroid visor during increased vertical pull in the laryngeal region" (page 202). We are currently developing an EMG study to further examine this issue.

In sum, the proposal that IF_0 is a deliberate enhancement of the speech signal seems somewhat tenuous. Even when F_0 is used extensively for other purposes and vowels occur with the whole range of F_0s, as in tone languages, IF_0 persists. It would seem that only a perceptual parsing of the F_0 signal could disentangle these effects, and this stage is later than the proposed enhancement effect, which is meant to be a low-level one. IF_0 disappears at the lower range of F_0s, a fact that is difficult to encompass in an enhancement account. Other enhancements of speech occasionally fail to be adopted by one language or another, while IF_0 seems to be universal. All of these facts together make it appear that IF_0 is truly "intrinsic" and not a deliberate enhancement of the speech signal.

4. Toward the source of IF_0

Our survey has provided some useful limitations on what a theory of IF_0 must accommodate. First is its universality. We found no evidence of unusual languages in our survey, and we have examined data from languages belonging to more than a third of the world's language families. Any theory that explains IF_0 must deal with the difference between the effect in the non-low portion of the F_0 range and the low. In the low region, IF_0 disappears. We believe that this is the difference between the changes in F_0 that can be accomplished via subglottal air pressure and CT activity versus the changes that need active lowering via the strap muscles. If the strap muscles completely counteract the effect of the tongue on the hyoid bone, then we would expect for low F_0 vowels to have no IF_0. Similar speculations appear in Pettorino (1987). Much more work remains to be done before this theory can be fully tested.

While our survey has not revealed any language that seems to have an especially exaggerated IF_0 effect, there are some articulatory sources for an exaggerated effect. In one experimental paradigm, if a speaker's jaw is fixed to a more open position than is normal for a vowel, then she will compensate with an exaggerated tongue movement. A stronger pull of the tongue should result in a larger IF_0 difference. This is what was found in an experiment by Ohala & Eukel (1987). While this experiment does not fully determine the mechanism involved, it does indicate that the action of the tongue is critically involved.

Another instance of an exaggerated effect is found in the speech of the deaf. Bush (1981) found larger IF_0 differences for deaf children compared with normal controls. Clearly, the source of this difference cannot be due to an auditorily based enhancement of the height difference. Part of the effect may have been the use of a higher F_0 (which may by itself increase IF_0; see above). But there seems to be an element of exaggerated articulation that by itself also increases the IF_0 effect. Perkell, Lane, Svirsky & Webster (1992) also found an exaggerated difference for subjects prior to receiving a cochlear implant, and this exaggeration disappeared after the implantation. The finding has not been universal, however. Lane & Webster (1991), studying some of the same subjects as Perkell *et al.* (1992), found no evidence of IF_0 at all in their deaf subjects. Lane & Webster measured productions of a full text (the Rainbow Passage), while Perkell *et al.* examined keywords in a carrier sentence. Perhaps the greater variability in the text production washed out any vowel effect that there might have been. But it does seem that certain forms of production can affect the IF_0.

There does not appear to be a developmental trend in IF_0. In another paper (Whalen, Levitt, Hsiao, & Smorodinsky, 1995), we show that even infants babbling at 6 months show IF_0. While it is true that, since every language has IF_0, infants could be imitating IF_0, it is hard to see how they could do this intentionally, since the infants do not have any vowel categories to enhance. Similarly, the data of Peterson & Barney (1952), Peterson (1961), Sorenson (1989), Glaze, Bless & Susser (1990) and Hillenbrand *et al.* (1995) also do not show any change in the size of the effect in the range from 5 to 11 years of age. If enhancements are something to be learned after the major components of a category are mastered, then the case for enhancement status of IF_0 is again weakened.

The difference between males and females is in need of verification and an explanation. Since many of the studies examined here involved only one sex or the other, it is difficult to fully accept the difference that appeared. If it holds up in further studies, it seems that the changes in male voices at puberty is likely to be responsible.

IF_0 appears to be universal. Our sampling of 31 languages, while far short of the 6000 or so total languages, covers a fairly wide range of families (11 of 29) and language types (tone, pitch accent, and stress). If languages that have no IF_0 constitute some small percentage of languages, then we might yet discover a language that makes no use of IF_0 at all. Although such propositions can never be fully tested, we feel that the results so far justify the assumption that IF_0 is universal. We can, then, for the moment, conclude that IF_0 is not a deliberate enhancement of the signal but rather a direct result of vowel articulation.

This project was supported by NIH grants DC-00403 and HD-01994 to Haskins Laboratories. Ken de Jong contributed a significant amount of work in extracting the Navaho F_0s, for which we are extremely grateful. Julie Lavoie also kindly provided her raw data. We are also grateful to the respondents from the LINGUIST list: Jan-Olaf Svantesson, Jurika Bakran, Ocke Bohn, Michael Jessen, Hartmut Traunmüller, Conrad Ouellon, Florien Koopmans-van Beinum, and Ingegerd Eklund. John Ohala and Harmut Traunmüller were especially helpful in ferreting out ever more references. We thank Carol A. Fowler, Ilse Lehiste, Kiyoshi Honda and John Ohala for helpful comments.

References

Antoniadis, Z. & Strube, H. W. (1981) Untersuchungen zum ⟨⟨intrinsic pitch⟩⟩ deutscher Vokale, *Phonetica*, **38**, 277–290.

Atkinson, J. (1973) *Aspects of intonation in speech: Implications from an experimental study of fundamental frequency.* Unpublished Ph.D. thesis, University of Connecticut, Storrs.

Autesserre, D., Roubeau, R., DiCristo, A., Chevrie-Muller, C., Hirst, D., Lacau, J. & Maton, B. (1987) Contribution du cricothyroidien et des muscles sous-hyoidiens aux variations de la frequence fondamentale en français: approche eletromyographique. In *Proceedings XIth International Congress of Phonetic Science*, **3** (pp. 35–38). Tallin, Estonia: Academy of Sciences of the Estonian SSR.

Bakran, J. & Stamenkovic, M. (1990) Inherentna frekvencija larngalnog tona u Hrvatskom standardnom jeziku, *Govor*, **7**, 1–20.

Black, J. W. (1949) Natural frequency, duration and intensity of vowels in reading, *Journal of Speech and Hearing Disorders*, **14**, 216–221.

Bolla, K. (1981) *A conspectus of Russian speech sounds.* Köln: Böhlau.

Bush, M. (1981) *Vowel articulation and laryngeal control in the speech of the deaf.* Ph.D. thesis, M.I.T.

Crandall, I. B. (1925) The sounds of speech, *Bell System Technical Journal*, **4**, 586–626.

Crystal, D. (1987) *The Cambridge encyclopedia of language.* Cambridge: Cambridge University Press.

Dave, R. (1967) A formant analysis of the clear, nasalized, and murmured vowels in Gujarati, *Indian Linguistics*, **28**, 1–30.

de Jong, K. & McDonough, J. (1993) Tone and tonogenesis in Navajo, *UCLA Working Papers in Phonetics*, **84**, 165–182.

DiCristo, A. (1982) *Prolegomenes a l'etude de l'intonation: micromelodie.* Paris: Editions du Centre National de La Recherche Scientifique.

DiCristo, A., Hirst, D. J. & Nishinuma, Y. (1979) L'estimation de la F_0 intrinsèque des voyelles: etude comparative, *Travaux de L'Institut de Phonétique D'Aix-en-Provence*, **6**, 149–176.

Diehl, R. L. (1991) The role of phonetics within the study of language, *Phonetica*, **48**, 120–134.

Diehl, R. L. & Kluender, K. R. (1989a) On the objects of speech perception, *Ecological Psychology*, **1**, 121–144.

Diehl, R. L. & Kluender, K. R. (1989b) Reply to commentators, *Ecological Psychology*, **1**, 195–225.

Fant, G. (1959) Acoustic analysis and synthesis of speech with applications to Swedish, *Ericsson Technics*, **1**, 3–108.

Fischer-Jørgensen, E. (1990) Intrinsic F_0 in tense and lax vowels with special reference to German, *Phonetica*, **47**, 99–140.

Flanagan, J. L. (1955) A difference limen for vowel formant frequency, *Journal of the Acoustical Society of America*, **27**, 613–617.
Fox, R. A. (1982) Individual variation in the perception of vowels: implications for a perception-production link, *Phonetica*, **39**, 1–22.
Gandour, J. & Maddieson, I. (1976) Measuring larynx movement in Standard Thai using the cricothyrometer, *Phonetica*, **33**, 241–267.
Gandour, J. & Weinberg, B. (1980). On the relationship between vowel height and fundamental frequency: evidence from esophageal speech, *Phonetica*, **37**, 344–354.
Glaze, L. E., Bless, D. M. & Susser, R. D. (1990) Acoustic analysis of vowel and loudness differences in children's voice, *Journal of Voice*, **4**, 37–44.
Hallé, P. A. (1994) Evidence for tone-specific activity of the sternohyoid muscle in modern standard Chinese, *Language and Speech*, **37**, 103–123.
Han, M. S. (1969) *Vietnamese tones.* Studies in the Phonology of Asian Languages, vol. 8. Los Angeles: University of Southern California.
Han, M. S. & Weitzman, R. S. (1967) *Acoustic features in the manner-differentiation of Korean stop consonants.* Studies in the Phonology of Asian Languages, vol. 5. Los Angeles: University of Southern California.
Higgins, M. B., Netsell, R. & Schulte, L. (1994) Aerodynamic and electroglottographic measures of normal voice production: intrasubject variability within and across sessions, *Journal of Speech and Hearing Research*, **37**, 38–45.
Hillenbrand, J., Getty, L. A., Clark, M. J. & Wheeler, K. (1995) Acoustic characteristics of American English vowels, *Journal of the Acoustical Society of America*, **97**, 3099–3111.
Hombert, J. M. (1977) Consonant types, vowel height and tone in Yoruba, *Studies in African Linguistics*, **8**, 173–190.
Homma, Y. (1973) An acoustic study of Japanese vowels, *Study of Sounds*, **16**, 347–368.
Honda, K. & Fujimura, O. (1991) Intrinsic vowel F_0 and phrase-final F_0 lowering: phonological vs. biological explanations. In *Vocal fold physiology: acoustic, perceptual, and physiological aspects of voice mechainsms* (J. Gauffin & B. Hammarbergs editor), pp. 149–157. San Diego, CA: Singular Publishing Group.
House, A. S. & Fairbanks, G. (1953) The influence of consonant environment upon the secondary acoustical characteristics of vowels, *Journal of the Acoustical Society of America*, **25**, 105–113.
Iivonen, A. K. (1989) Regionally determined realization of the standard German vowel system, *Mimeographed Series of the Department of Phonetics, University of Helsinki*, **15**, 21–28.
Ivic, P. & Lehiste, I. (1965) Prilozi ispitivanju fonetske i fonoloske prirode akcenata u savremenom srpskohravatskom jeziku: 2, *Zbornik za Filologiji i Lingvistiku*, **8**, 75–117.
Kim, K. (1968) F_0 variations according to consonantal environments. Phonology Laboratory, University of California, Berkeley.
King, L., Ramming, H., Schiefer, L. & Tillmann, H. G. (1987) Initial F_0-contours in Shanghai CV-syllables—an interactive function of tone, vowel height, and place and manner of stop articulation. In *Proceedings XIth International Congress of Phonetic Sciences*, **1** (pp. 154–157). Tallinn, Estonia: Academy of Sciences of the Estonian SSR.
Kingston, J. (1993) The phonetics and phonology of perceptually motivated articulatory covariation, *Language and Speech*, **35**, 99–113.
Koopmans-van Beinum, F. J. (1980) *Vowel contrast reduction: an acoustic and perceptual study of Dutch vowels in various speech conditions.* Amsterdam: Academische Pers.
Ladd, D. R. & Silverman, K. E. A. (1984) Vowel intrinsic pitch in connected speech, *Phonetica*, **41**, 31–40.
Ladefoged, P. (1968) *A phonetic study of West African lnaguages: an auditory-instrumental study.* Cambridge: Cambridge University Press.
Ladefoged, P. & Maddieson, I. (1990) Vowels of the world's languages, *Journal of Phonetics*, **18**, 93–122.
Lane, H. & Webster, J. (1991) Speech deterioration in postlingually deafened adults, *Journal of the Acoustical Society of America*, **89**, 859–866.
Lavoie, J. (1994) La fréquence intrinsèque des voyelles en français québécois. In *Actes des huitsieme Journées de linguistique* (1994) (F. Z. Belyazid, S. Belyazid, G. Cochrane, J. Côté, J. de Blois, M. Faucher, F. Jean & W. Zoualis, editors) pp. 109–113. Quebec City: Université Laval.
Lehiste, I. & Peterson, G. E. (1961) Some basic considerations in the analysis of intonation, *Journal of the Acoustical Society of America*, **33**, 419–423.
Lindblom, B. (1986) Phonetic universals in vowel systems. In *Experimental phonology* (J. J. Ohala & J. J. Jaegers editors), pp. 13–44. Orlando, FL: Academic.
Lounsbury, F. (1953) *Oneida verb morphology*, New Haven, CT: Yale University Press.
Maddieson, I. (1984) *Patterns of sounds.* Cambridge: Cambridge University Press.
Meyer, E. A. (1896–7) Zur Tonbewegung des Vokals im gesprochenen und gesungenen einzelwort, *Phonetische Studien* (*Beiblatt zu der Zeitschrift Die Neuren Sprachen*), **10**, 1–21.

Möbius, B., Zimmermann, A. & Hess, W. (1987) Microprosodic fundamental frequency variations in German. In *Proceedings XIth International Congress of Phonetic Sciences*, **1** (pp. 146–149). Tallinn, Estonia: Academy of Sciences of the Estonian SSR.

Mohr, B. (1971) Intrinsic variations in the speech signal, *Phonetica*, **23**, 65–93.

Nearey, T. M. (1989) Static, dynamic, and relational properties in vowel perception, *Journal of the Acoustical Society of America*, **85**, 2088–2113.

Neweklowsky, G. (1975) Spezifische Dauer und spezifische Tonhöhe der Vokale, *Phonetica*, **32**, 38–60.

Nishinuma, Y. (1979) *Un modèle d'analyse automatique de la prosodie: accent et intonation en japonais.* Paris: Centre National de la Recherche Scientifique.

Nittrouer, S., McGowan, R. S., Milenkowic, P. H. & Beehler, D. (1990) Acoustic measurements of men's and women's voices: a study of context effects and covariation, *Journal of Speech and Hearing Research*, **33**, 761–775.

Ohala, J. J. & Eukel, B. W. (1987) Explaining the intrinsic pitch of vowels. In *In honor of Ilse Lehiste* (R. Channon & L. Shockey, editors), pp. 207–215. Dordrecht: Foris.

Pakerys, A. (1982) *Lietuviu bendrines kalbos prozodija.* Vilnius: Mokslas.

Perkell, J., Lane, H., Svirsky, M. & Webster, J. (1992) Speech of cochlear implant patients: A longitudinal study of vowel production, *Journal of the Acoustical Society of America*, **91**, 2961–2978.

Peterson, G. E. (1961) Parameters of vowel quality, *Journal of Speech and Hearing Research*, **4**, 10–29.

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

Pettorino, M. (1987) Intrinsic pitch of vowels: an experimental study on Italian. In *Proceedings XIth International Congress of Phonetic Sciences*, **1** (pp. 138–141). Tallinn, Estonia: Academy of Sciences of the Estonian SSR.

Pilszczikowa-Chodak, N. (1972) Tone-vowel height correlation and tone assignment in the patterns of verb and noun plurals in Hausa, *Studies in African Linguistics*, **3**, 399–421.

Rakotofiringa, H. (1968) *Contributions a l'étude de la phonétique malgache II: hauteur, durée et intensité vocaliques efficaces.* Université de Grenoble.

Rakotofiringa, H. (1982) *L'accent et les unites phoniques elementaires de base en malgache-merina.* Lille: Atelier National de Reproduction des Theses, Université de Lille.

Reinholt Petersen, N. (1978) Intrinsic fundamental frequency of Danish vowels, *Journal of Phonetics*, **6**, 177–189.

Rossi, M. & Autesserre, D. (1981) Movements of the hyoid and the larynx and the intrinsic frequency of vowels, *Journal of Phonetics*, **9**, 233–249.

Samaras, M. (1972) Influence de l'entourage consonantique sur les variations de la frequence laryngienne des voyelles du grec moderne, *Bulletin de l'Institut de Phonétique de Grenoble*, **1**, 57–66.

Sapir, S. (1989) The intrinsic pitch of vowels: Theoretical, physiological and clinical considerations, *Journal of Voice*, **3**, 44–51.

Schiefer, L. (1987) F_0 perturbations in Hindi. In *Proceedings XIth International Congress of Phonetic Sciences*, **1** (pp. 150–153). Tallinn, Estonia: Academy of Sciences of the Estonian SSR.

Shadle, C. H. (1985) Intrinsic fundamental frequency of vowels in sentence context, *Journal of the Acoustical Society of America*, **78**, 1562–1567.

Shi, B. & Zhang, J. (1987) Vowel intrinsic pitch in standard Chinese. In *Proceedings XIth International Congress of Phonetic Sciences*, **1** (pp. 142–145). Tallinn, Estonia: Academy of Sciences of the Estonian SSR.

Silvermann, K. E. A. (1987) *The structure and processing of fundamental frequency contours.* Unpublished Ph.D. thesis, University of Cambridge.

Sorenson, D. N. (1989) A fundamental frequency investigation of children ages 6–10 years old, *Journal of Communicative Disorders*, **22**, 115–123.

Steele, S. A. (1986) Interaction of vowel F_0 and prosody, *Phonetica*, **43**, 92–105.

Steffen-Batóg, M. (1970) The influence of intrinsic vowel pitch on the differences in the realizations of intended intervals, *Speech Analysis and Synthesis*, **2**, 177–194.

Svantesson, J. (1988) Voiceless stops and F_0 in Kammu, *Lund University Department of Linguistics Working Papers*, **34**, 116–119.

Svantesson, J. (1993) Phonetic correlates of register in Paraok, *Reports from Uppsala University Linguistics* (RUUL), **23**, 102–105.

Tamas, S. (1976) *A beszédfolyamat alaptényezói.* Budapest: Akadémiai Kiadó.

Taylor, H. C. (1933) The fundamental pitch of English vowels, *Journal of Experimental Psychology*, **16**, 565–582.

Traunmüller, H. (1982) Der vokalismus im ostmittelbairischen, *Zeitschrift für Dialektologie und Linguistik*, **3**, 289–333.

van Son, R. (1993) *Spectro-temporal features of vowel segments.* Amsterdam: IFOTT.

Vilkman, E., Aaltonen, O., Raimo, I., Arajärvi, P. & Oksanen, H. (1989) Articulatory hyoid-laryngeal

changes *vs*. cricothyroid muscle activity in the control of intrinsic F_0 of vowels, *Journal of Phonetics*, **17**, 193–203.
Whalen, D. H., Levitt, A. G., Hsiao, P.-L. & Smorodinsky, I. (1995) Intrinsic F_0 of vowels in the babbling of 6-, 9- and 12-month-old French- and English-learning infants, *Journal of the Acoustical Society of America* **97**, 2533–2539.
Zawadzski, P. A. & Gilbert, H. R. (1989) Vowel fundamental frequency and articulatory position, *Journal of Phonetics*, **17**, 159–166.
Zee, E. (1980) Tone and vowel quality, *Journal of Phonetics*, **8**, 247–258.
Zhu, X. (1994) *Shanghai tonetics.* Unpublished Ph.D. dissertation, The Australian National University, Canberra, NSW.