The frequency range of the voice fundamental in the speech of male and female adults

Hartmut Traunmüller and Anders Eriksson

Institutionen för lingvistik, Stockholms universitet, S-106 91 Stockholm, Sweden.

Published data on the frequency of the voice fundamental (F_0) in speech show its range of variation, often expressed in terms of two standard deviations (SD) of the F_0-distribution, to be approximately the same for men and women if expressed in semitones, but the observed SD varies substantially between different investigations. Most of the differences can be attributed to the following factors: SD is increased in tone languages and it varies with the type of discourse. The more ‘lively’ the type of discourse, the larger it is. The dependence of SD on the type of discourse tends to be mom pronounced in the speech of women than of men. Based on an analysis of various production data A is shown that speakers normally achieve an increased SD by increasing the excursions of F_0 from a ‘base-value’ that lies about 1.5 SD below their mean F_0. This is relevant to applications in speech technology as well as to general theories of speech communication such as the ‘modulation theory’ in which the base-value of F_0 is seen as a carrier frequency.

INTRODUCTION

There is a substantial amount of data on the frequency of the voice fundamental (F_0) in the speech of speakers who differ in age and sex. Such data have been published for several languages and for various types of discourse. The data reported nearly always include an average measure of F_0, usually expressed in Hz, but in some cases the average duration of a period has been reported instead. Typical values obtained for F_0 are 120 Hz for men and 210 Hz for women. The mean values change slightly with age. For men, the decrease in F_0 that is most dramatic during puberty has been observed to continue with successive deceleration until about 35 years of age At bout 55 years of age, F_0 begins to rise again (Hollien and Ship, 1972; Kitzing, 1979; Pegoraro-Krook, 1988). For women, F_0 is stationary up to the age of menopause, when it decreases to reach a minimum that is about 15 Hz lower around 70 years of age (Chevrie-Muller et al., 1971; Kitzing, 1979; Stoicheff, 1981; Pegoraro-Krook, 1988). The physiological changes responsible for this can be understood as an effect of the increased testosterone-oestrogen ratio. A similar lowering of F_0 can be induced by the habit of smoking (Gilbert and Weismer, 1974).

Most studies also report on the between-speaker spread in average F_0 for each sex. The present paper, however, is primarily concerned with the description of the within-speaker variations in F_0 of male and female adults. Statistical data on the distribution of F_0-values or on the F_0-range used by each speaker have been included in quite a large number of studies. Some of these have shown that the F_0-range is influenced by various factors such as the language, the type of text, the type of discourse, and the emotional state of the speaker. We want to gain an overview of the effects of these factors on the F0-range used and we want to know in which way speakers expand and contract their F_0-range, as far as this can be described in a general way.

The present evaluation of published data has been initiated in preparation of an experimental investigation of the perception of F_0-excursions (Traunmüller and Eriksson, 1994). The method used in that study involved simulations of various para- and extralinguistic variations, including speaker sex, in addition to variations in the extent of the F_0-excursions.

MEAN VALUES AND STANDARD DEVIATIONS OF F_0 REPORTED IN THE LITERATURE.

Unfortunately, the statistics of F_0-values are often not very well described by a normal distribution. If F_0 is scaled linearly (in Hz), there is, typically, some positive skewness. Analysis of the duration of periods revealed an even stronger skewness (Mikeev, 1971). In addition, it has been observed that some speakers show a bimodal F_0-distribution, in particular when speaking with increased vocal effort, as in a parliamentary debate (Rappaport, 1958). In order to compare the results from studies in which different ways of describing the F_0-variation have been chosen, we are forced to assume normality. We will, however, not include any reports for which this assumption appears to involve a risk of introducing a substantial error. The results of some of the remaining studies are summarized in Table I. The table includes only those investigations in which both male and female adult speakers performed the same kind of task. A considerable number of investigations, mainly on English, had to be left aside since they involved only one sex or the use of incommensurate statistics.

The original reports summarized in Table I contain data on average F_0 and on the average standard deviation (SD) of F_0 per speaker reported in Hz, in semitones, or as a frequency modulation factor (SD/mean) in %. In some cases, the range was reported in terms of two SD in semitones. In all except one of the reports (Rose, 1991), the average F_0 was clearly higher and the F_0-range in Hz clearly wider for women than for men. The between-sex difference more or less disappears for F_0-range if it is expressed in semitones or as a modulation factor. Pooling all the results listed in Table I does not reveal any significant sex difference in SD expressed in semitones, but if SD is expressed in Hz, in barks, or in ERB (Moore and Glasberg, 1983), the value obtained for women is higher than that for men in all the cases listed, except for Rose (1991).

The very high values of average F_0 observed in male speakers of Wú dialects of Chinese (Rose, 1991) are quite remarkable. They show that even the average F_0 used in speech belongs to the set of properties that can be prescribed by social convention. Although these Chinese dialects present an extreme case, the phenomenon is not unique. An increased average F_0 can also be observed in the Swedish dialect spoken in the province of Småland (Elert and Hammarberg, 1991). In most languages, however, the F_0-range used by speakers appears to be governed by the principle of minimizing the physiological effort. Speakers comply with this principle by using the lower part of their physiological F_0-range. The lowest F_0 a speaker uses in ordinary speech is often the same as the lowest F_0 at which he is capable of sustaining phonation, while the upper limit, that can only be reached with extremely tensed vocal folds, is only approached in exceptional cases such as when shouting for help in an emergency situation. Voice range profiles (phonetograms) which show the lowest and the highest F_0 at which a speaker is capable of sustaining phonation as a function of sound pressure level (SPL), such as recorded by Pabon and Plomp (1988), show the absolute minimum values of F_0, in fact, to be slightly higher than the minimum values of F_0 observed for male speakers in most of the investigations listed in Table 1. Phonetograms show F_{0min} to rise with SPL and in unrestrained speech F_0 has also been observed to increase with an increase in vocal effort (Ladefoged, 1967; Rostolland, 1982). An increase in muscular tonus caused by emotional factors can also lead to an increase in F_{0min}. For a detailed physiological model see Titze (1989).

Table I. Mean value of F_0 in Hz and average F_0-variation (SD) in semitones according to ten investigations that report results from adult male and female speakers in the same setting. Under 'Type', the speech samples are classified according to their expected liveliness, as explained in text.

Investigation	**Type**	**n**	**Sex**	**Age**	$\mathbf{F_0}$	**SD**
Rappaport (1958), Geman	1	190	m		129	2.3
	1	108	f		238	1.9
Chevrie-Muller *et al*. (1967), French	2	21	m	20–61	145	2.5
	2	21	f	19–72	226	2.3
Takefuta *et al.* (1972), English	4	24	m		127	3.8
	4	24	f		186	5.4
Chen (1974), Mandarin Chinese	2	2	m	30–50	108	4.1
	2	2	f	30–50	184	3.8
Boë *et al.* (1975), French	2	30	m		118	18
	2	30	f		207	3.0
Kitzing (1979), Swedish	2	51	in	21–70	110	3.0
	2	141	f	21–70	193	2.7
Johns-Lewis (1986), English:						
Conversation	2	5	m	24–49	101	3.4
	2	5	f	24–49	182	2.7
Reading	3	5	m	24–49	128	4.35
	3	5	f	24–49	213	4.5
Acting	4	5	m	24–49	142	4.85
	4	5	f	24–49	239	5.3
Graddol (1986), English:						
Reading passage A	1	12	m	25–40	119	3.6
	3	15	f	25–40	207	3.05
Reading passage B	3	12	m	25–40	131	4.55
	3	15	f	25–40	219	3.9
Pegoraro Krook (1988), Swedish	2	198	m	20–79	113	2.65
	2	467	f	20–89	188	2.55
Rose (1991), Wú	2	4	m	25–62	170	4.1
	2	3	f	30–64	187	3.8
Average/investigation		11	m		124	3.4
European languages only		11	f		211	3.4
Average/balanced speaker		471	m		119	2.8
European languages only		471	f		207	2.7

As for the extent of F_0-excursions, it is known that these are influenced by conventional linguistic factors reflected in the language and text in question and by various paralinguistic factors. In linguistic terms, the extent of the F_0-excursions in an utterance can be referred to as its “prosodic explicitness”.

Locally, the explicitness of the prosody within an utterance is affected by the placement of focal and contrastive stress. More globally, the extent of F_0-excursions is affected by attitudinal and emotional factors. Emotionally depressed, sad or ashamed speakers produce speech with very little variation in F_0, while increased variation in F_0 reflects an excited emotional state in the speaker, such as surprise, interest, and joy, but also contempt and anger (Fairbanks and Pronovost, 1939; Fónagy and Magdics, 1963; Williams and Stevens, 1972; Scherer, 1974; van Bezooyen, 1984). Increased F_0-excursions can also be observed in speech directed to infants (Garnica, 1977). Instead of reflecting an emotionally excited state of the *speaker,* in this case, the increased F_0-excursions appear to serve the purpose of evoking and maintaining a positively excited emotional state in the *listener*. Ohala (1983) suggested that The phenomenon might be understood as one of the manifestations of the cross species association of high pitch vocalizations with lack of threat, but since the increased average F_0 in this case is mainly due to an increased F_0-range, this appears unlikely to be the whole explanation.

As for the linguistic factor, we would expect F_0-excursions to be more frequent and, due to the need of additional distinctiveness, probably also larger in tone languages than in languages that do not use tone for segmental distinctions. This has been confirmed in a comparison of Northern Chinese and English (Chen, 1974) where it is also shown that speakers of English with Chinese as a second language use more extensive F_0-excursions in their Chinese than in their English, but that native speakers of Chinese use still more extensive F_0-excursions. Between languages, we would expect the effect of tone on the extent of the F_0-excursions to increase with the complexity of the tone system. In tone languages with only two distinctive level tones, the need of additional distinctiveness would seem to be more moderate than in languages with more complex tone systems including ‘contour’ tones.

The extent of F_0-excursions in speech increases slightly with age. Chevrie-Muller et al. (1971) reported a slight but significant increase with age ($p < 0.01$) in a study including 104 female speakers between 13 and 99 years A similar increase can be seen in the data by Pegoraro-Krook (1988) if the highest age-decade is excluded, although this went unnoticed. A test based on the published values showed a significant increase in SD (at p < 0.0 1) from 20 to 79 years for the female as well as for the male subjects.

Although not included in Table 1, it is relevant to add that the habit of smoking has not only a lowering affect on the mean F_0 in females, but it has also the effect of increasing their F_0-range quite substantially. The data obtained by Gilbert and Weismer (1974) imply an SD of 3.25 semitones for smokers (F_{0mean} = 164 Hz) to compare with 2.35 semitones for non-smokers (F_{0mean} = 183 Hz).

Expressed in semitones or as a percentage of F_0, the between speaker variation in average F_0 is generally reported to be higher for men than for women. Although this is
outside the focus of the present paper, we would like to propose an explanation for why this should be so. Since in most investigations the speakers have not been checked for smoking, it might be suspected that the larger variation among men may possibly be due to a larger number of smokers among the male speakers. However, a similar result should be expected even if smokers

were excluded: The between-speaker variance in average F_0 can be considered to consist of two components, one that is present in men and women to the same extent, and one that is specific for men. The common component is analogous to individual variation in body size. For men, however, an additional amount of variance is added as a consequence of individual variations in the laryngeal changes occurring during puberty, since these reflect the impact of hormone levels which are likely to vary between individuals. Stoicheff's (1981) finding that the between-speaker variation in average F_0 in the speech of (non-smoking) women increases after menopause also fits into this picture.

MEAN F_0, SD, AND DISCOURSE TYPE

When reading aloud, it has been shown that the type of text has a significant effect on the SD of F_0 (Graddol, 1986), but the effects on SD of variations in the type of discourse such as 'conversation' compared with 'acting' are larger (Johns-Lewis, 1986).

Based on the descriptions of the various types of speech material which resulted in the data summarized in Table I, we have estimated the degree of liveliness that might be expected in the type of discourse used in each case. This has been done by assigning one of four liveliness classes to each type of discourse. The business conversations by telephone, analyzed by Rappaport (1958) we have put into the lowest liveliness class. The second class contains somewhat more personal conversations and such tasks as reading a text for the purpose of clinical investigation of one's voice. The third class contains cases where texts have been read aloud in such a way that it can be assumed that the subjects attempted to read in a pleasant way. Into the highest class we have put Johns-Lewis' 'acting' and the investigation by Takefuta *et al*. (1972), who had asked their subjects to "produce as many different intonations as they could think of' when repeatedly producing a set of given sentences of the kind that can easily be loaded with various paralinguistic meanings.

Table II. Average F_0-variation (SD in semitones) as a function of the type of speech as classified in Table I, sexes pooled. For each investigation in which the SD was higher for women than for men, a "+" sign is shown. In contrary cases, a '-' sign has been entered.

	European lang.		Chinese lang.	
Liveliness class	SD	N	SD	N
(4) Very high	4.8	+ +		
(3) High	4.0	+ – –		
(2) Moderate	2.8	– + – – –	4.0	– –
(1) Low	2.1	–		

For each liveliness class we have calculated the average SD (in semitones) keeping the tone languages apart from the rest. The result is shown in Table II. Although the liveliness classification is somewhat arbitrary, the table can be said to illustrate the following four points:

1) The SD of F_0 increases with increasing 'liveliness' of the discourse.

2) The SD of F_0 is larger in tone languages than in non-tone languages.

3) The SD of F_0 is approximately the same in the speech of men and women if expressed in semitones or as a frequency modulation factor.

4) In the most lively types of discourse, women show a larger SD than men, while their SD (in semitones) tends to be lower than that of men in the least lively types of discourse.

While the first three points do not call for further discussion, the fourth point should be regarded as tentative. The apparently larger elasticity of the F_0 range in the speech of women can be seen in the investigation by Johns-Lewis (1986), in which the mean SD in acting was higher than that in conversation by 1.45 semitones for men and by 2.6 semitones for women. If we compare the extreme values among all the investigations included in Table I, we obtain a difference of 2.55 semitones for men and 3.5 semitones for women. However, the large elasticity of female F_0 shows itself mainly in ‘acting’ (class 4) and if this class were removed, the data would no longer suggest any sex-difference. In the other types of discourse, women show mostly a lower value of SD than men. However, in the large study by Pegoraro-Krook (1988), this holds only for the age groups 50–59 and 60–69 (123 male, 168 female subjects). For each of the age groups 20–29, 30–39, 40-49, and 70-79 (75 male, 281 female subjects), the SD was higher in the speech of women as compared with men. If these age groups had been shown separately in Table II, there would be five ‘+’ signs as well as five ‘–’ signs for class 2 in the second column.

In order to account for the substantial sex-difference in class 4, it might be suggested that men do not need to produce as large Fo-excursions as women do in order to evoke the same perceived degree of liveliness, but this hypothesis has to be rejected on the basis of the results of the investigation of the perception of liveliness (Traunmüller and Eriksson, 1994). Instead, it appears likely that the difference is due to physiological, or possibly psychological or cultural factors. It may, for example, be the case that a smooth transition into the falsetto register is more difficult to achieve for men than for women or that men, perhaps as an adaptation to such a physiological factor, consider a lower degree of liveliness to be appropriate.

TRANSFORMATIONS of F_0-CONTOURS

If it is the case that the lowest F_0 frequency speakers use in an utterance is near the floor of their physiological F_0-range then we should expect the mean value of F_0 to increase with increasing SD. This is confirmed by the data of Johns-Lewis (1986) and Graddol (1986), listed in Table I. These data also allow us to find out more precisely how the expansion of the F_0-excursions is performed when a speaker increases his liveliness, *ceteris paribus.* They allow us to calculate a ‘base-value’ of F_0, that remains invariant when a speaker varies the liveliness of his speech. The data must be interpreted with some caution since the texts used in the different types of discourse were not the same and we cannot be sure that the speakers did not vary any additional property of their speech, such as vocal effort, that would affect F_0.

There is, however, an investigation by Bruce (1982) in which an actress was asked to produce sentences first with a detached and then with an involved attitude, so that our ‘ceteris paribus’ condition is satisfied.

In Bruce’s (1982) study, the F_0-values of the local minima and maxima of the F_0-contour were reported. Fig. 1 shows, for each minimum and maximum, the excess of the F_0-value in the involved version over that of the corresponding point on the F_0-contour of the detached version (in semitones) as a function of the F_0-value in the detached version. The regression line in Fig. I describes these data fairly well, i.e., it explains 76% of the variance. The F_0-value corresponding to the point where the regression line crosses the horizontal zero-line is F_b, the invariant base-value of F_0 we are looking for. If the F_0-distribution is normal, the frequency position of F_b can be calculated as

$$F_b = F_{mean} - k \cdot \sigma (F) \quad (1)$$

Since this should hold for any value of σ, it is possible to obtain an estimate of F_b even on the basis of one single utterance, given that k is known. Although in Fig. 1 a logarithmic scaling of pitch has been chosen, the choice of scale is actually not very crucial in this case. Linear regression lines fit the data equally well if a linear (Hz), tonotopic (bark), equivalent rectangular bandwidth (ERB), or logarithmic (semitones) scale of pitch is used.

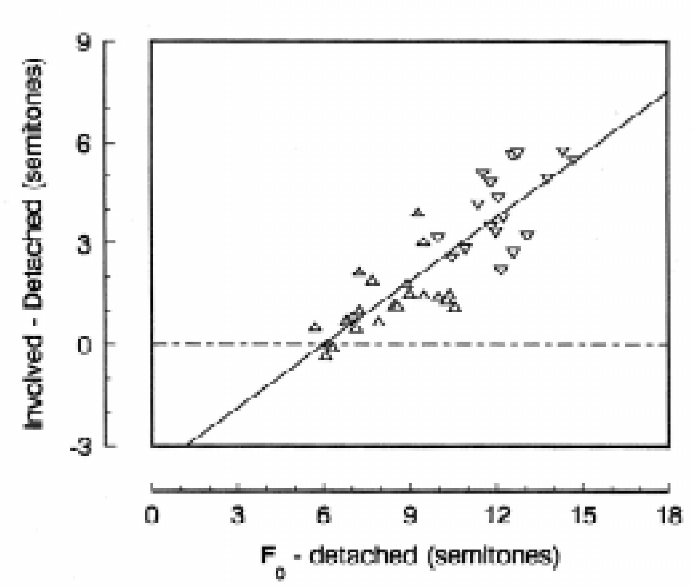

Figure 1. Maxima and minima in the F_0-contour of four utterances produced with a detached and an involved attitude by a female speaker of Swedish. Mean values from six repetitions. F_0-excess in involved version plotted against F_0-values in detached version. Regression line also shown (r = 0.86). Data from Bruce (1982).

Fig. 2 shows the F_0-data for each of 5 male and 5 female speakers in three types of discourse: conversation, reading aloud, and acting. These are the data obtained by Johns-Lewis (1986). The majority of the speakers, 3 male and 4 female, showed a uniform behaviour: Average F_0 and F_0-range (SD) have the smallest values in conversation. Both values are higher in reading aloud, and highest in acting. Except for the between-speaker differences in mean F_0, none of these speakers deviated much from the average shown by the dashed line. The remaining 3 speakers, 2 male and 1 female, showed, at some point, a change in F_0 without change in F_0-range. This may be due to a modification in vocal effort instead of prosodic explicitness. The data from the other 7 speakers show no effects in addition to that of having adapted their prosodic explicitness to the type of discourse. As distinct from the case shown in Fig. 1, the choice of scaling is crucial here. Due to the between-speaker variation in average F_0, Fig. 2 would look different if F_0 had not been scaled in semitones and our conclusion that the majority of speakers behaved in a uniform way would retain its validity only in a qualitative sense.

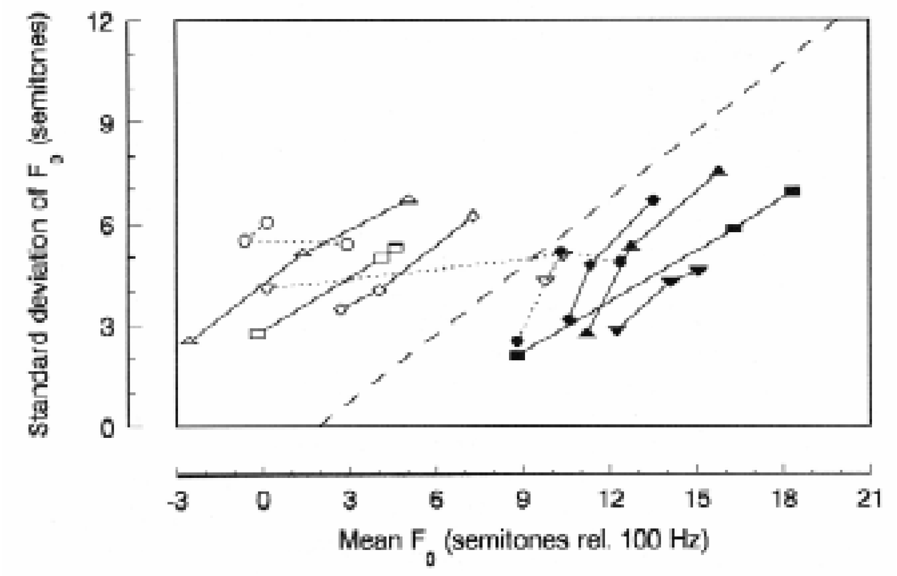

Figure 2. F_0 data of 5 male and 5 female speakers (open and filled symbols) in three types of discourse: conversation, reading aloud, and acting; connected by lines in this order. Data from Johns-Lewis (1986). Regression line (dashed) fitted to the average of the 7 subjects who behaved in a similar way.

On the basis of the line that shows the average of the 7 uniformly behaving speakers in Fig. 2 it is possible to calculate the value of k in Eq. 1. We obtain k = 1.5 for this case. The data shown in Fig. 1 do not allow a precise calculation of k since the data points shown do not represent an unbiased sample of F_0-values, hence σ is not known precisely, but a reasonable estimate would be $1.6 < k < 2.0$. An approximate value of k can also be calculated on the basis of Graddol's data (1986), which include a comparatively large number of speakers, 12 male and 15 female, but the difference in the extent of the F_0-excursions between the two types of discourse is not so large, and therefore the data are somewhat obscured by statistical noise. We obtain k = 1.7 for male and k = 1.1 for female speakers. Although the variation in Graddol's data is not primarily due to variation in liveliness, it is not unreasonable to assume that speakers manipulate their F_0-range in approximately the same way as long as no change in vocal effort, voice register, or emotional tension is involved. Given these restrictions, the F_b of a speaker can, as a rule of thumb, be expected to be about 1.5 σ below his average F_0 in any type of discourse. If F_0-values have a normal distribution, F_0 will be higher than F_b 93 % of the time.

This result is compatible with the previous observation that F_0 at the endpoint of sentences with a falling F_0-contour, where it is close to F_b, shows comparatively little within speaker variation (Menn and Boyce, 1982; Liberman and Pierrehumbert, 1984), but it is also compatible with Ladd's (1988) observation that this does not necessarily hold between different experimental sessions.

We are now in a position where we can simulate natural variations in the extent of a given speaker's F_0-excursions by means of subjecting his speech to LPC-analysis and to resynthesize it after recalculation of the successive F_0 values, as we intended to do in the investigation of the perception of F_0-excursions (Traunmüller and Eriksson, 1994). This can be combined with an adjustment in F_b and an holistic recalculation of the formant frequencies, such as suggested by Traunmüller (1988) in order to modify various additional para- and extralinguistic properties of the speech signal, including speaker sex. For this purpose, the value of F_0 (for each analysis frame) can be recalculated according to the equation

$$f' = k_b[F_b + k_e\,(f - F_b)] \qquad (2)$$

where f' is the recalculated value of F_0 for a given analysis frame, f is its original value, k_e is the 'excursion factor' by which the deviation of F_0 from F_b is multiplied, and k_b is the 'base-value factor' that describes the relation between the recalculated and original values of F_b. The values of F_b for the average per balanced speaker of the European languages listed in Table I are 93.4 Hz for the male and 163.8 Hz for the female group.

In accordance with these data, adult male speech can be transformed into adult female speech by choosing $k_b = 1.75$, provided that the formant frequencies are also appropriately modified. A recalculation in accordance with Eq. (2) can, of course, also be applied to formant based speech synthesis by rule or, in general terms, to any kind of synthesis that allows independent control of voice source (F_0) and filter properties.

While the present analysis of production data has shown that we have to choose $k_e = 1.00$ if we want to transform 'typical' male speech into 'typical' female speech, we would like to stress that no such analysis allows any conclusions about perception to be drawn. 'Typical' female speech might still be perceived as more lively or less lively than 'typical' male speech.

The base-value F_b plays a central role in the recently proposed modulation theory of speech communication (Traunmüller, 1994) which considers speech signals as the result of allowing conventional linguistic and paralinguistic gestures to modulate, in a complex way, a carrier signal that conveys the extralinguistic information about the speaker (age, sex, vocal effort, etc.). The carrier signal is thought of as a neutral vowel, phonated at F_b. In order to segregate the different types of information, the listener has to demodulate the speech signal. As for F_0, this implies that he has to estimate the value of F_b, to evaluate the deviations of F_0 from F_b, and to apply what is equivalent to an automatic gain control in order to recover the linguistic information carried by F_0 irrespective of holistic variations in the extent of the F_0-excursions.

ACKNOWLEDGMENT

This research has been supported, in part, by a grant from HSFR, the Swedish Council for Research in the Humanities and Social Sciences, within the frame of the Swedish Language Technology Program.

REFERENCES

Bezooyen, R. van (1984). *Characteristics and Recognizability of Vocal Expressions of Emotion* (Foris Publications, Dordrecht).

Boë, L-J., Contini, M., and Rakotofiringa, H. (1975). "Étude statistique de la fréquence laryngienne," Phonetica **32**, 1–23.

Bruce, G. (1982). "Developing the Swedish intonation model," in *Working Papers* **22** (Lund university, Department of linguistics), 51–116.

Chen, G-T. (1974). "The pitch range of English and Chinese speakers," Journal of Chinese Linguistics **2**, 159–171.

Chevrie-Muller, C., and Gremy, F. (1967). “Contribution a l’établissement de quelques constantes physiologiques de la voix parlée de l’adulte,” Journal Français d’Oto-Rhino-Laryngologie **XV1**, 149–154.

Chevrie-Muller, C., Salomon, D., and Ferrey, G. (1971). “Contribution a l’établissement de quelques constantes physiologiques de la voix parlée de la femme adolescente, adulte et age,” Journal Français d’Oto-Rhino-Laryngologie **XVI**, 433–455.

Elert, C-C., and Hamrnarberg, B. (1991). “Regional voice variation in Sweden,” in *Actes du XIIème Congres International des Sciences Phonétiques,* Vol. 4 (Université de Provence, Service des Publications, Aix-en-Provence), 418–420.

Fairbanks, G., and Pronovost, W. (1939). “An experimental study of the pitch characteristics of the voice during the expression of emotion,” Speech Monographs **6**, 85–105.

Fónagy, 1, and Magdics, K. (1963). “Emotional patterns in intonation and music,” Phonetica **16**, 293–326.

Garnica, O. K. (1977). “Some prosodic and paralinguistic features of speech to young children,” in *Talking to Children: Language Input and Acquisition,* edited by Catherine E. Snow and Charles A. Ferguson (Cambridge University Press), pp. 63–88.

Gilbert, H.R., and Weismer, G.G. (1974). “The effects of smoking on the speaking fundamental frequency of adult women,” J. Psycholing. Res. **3**, 225–231.

Graddol, D. (1986). “Discourse specific pitch behaviour,” in *Intonation in Discourse,* edited by Catherine Johns-Lewis (Croom Helm, London and Sidney), pp. 221–237.

Hollien, H., and Ship, T. (1972) “Speaking fundamental frequency and chronological age in males,” J. Speech Hear. Res. **15**, 155–159.

Johns-Lewis, C**.** (1986). “Prosodic differentiation of discourse modes,” in *Intonation in Discourse* edited by Catherine Johns-Lewis (Croom Helm, London and Sidney), pp. 199–219.

Kitzing, P**.** (1979). *Glottografisk frekvensindikering: En undersökningsmetod för mätning av röstläge och röstomfång samt framställning av röstfrekvensdistributionen* (Lund University, Malmö)

Ladd, D.R. (1988). “Declination “reset” and the hierarchical organization of utterances,” J. Acoust. Soc. Am. **84**, 530–544.

Ladefoged, P. (1967). “Stress and respiratory activity,” in *Three Areas of Experimental Phonetics,* edited by Peter Ladefoged (Oxford University Press, London), pp, 1–49.

Liberman, M., and Pierrehumbert, J**.** (1984) “Intonational invariance under changes in pitch range and length,” in *Language Sound Structure,* edited by Mark Aronoff and Richard T. Oehrle (MIT Press, Cambridge, Mass), pp. 157–233.

Menn, L., and Boyce, S. (1982). “Fundamental frequency and discourse structure,” Language and Speech **25**, 341–383.

Mikeev, Y. V. (1971). “Statistical distribution of the periods of the fundamental tone of Russian speech,” Sov. Phys. Acoust. **16**, 474–477.

Moore, B.C.J., and Glasberg, B.R. (1983). “Suggested formulae for calculating auditoryfilter bandwidths and excitation patterns,” J. Acoust. Soc. Am. **74**, 750–753.

Ohala, J.J. (1983). “Cross-Language use of pitch: An ethological view,” Phonetica **40**, 1–18.

Pabon, J. P. H., and Plomp, R. (1988). “Automatic phonetogram recording supplemented with acoustical voice-quality parameters,” J. Speech Hear. Res. **31**, 710–722.

Pegoraro-Krook, M. 1. (1988). “Speaking fundamental frequency characteristics of normal Swedish subjects obtained by glottal frequency analysis,” Folia Phoniatrica **40**, 82–90.

Rappaport, W. (1958). “Über Messungen der Tonhöhenverteilung in der deutschen Sprache,” Acustica **8**, 220–225.

Rose, P. (1991). “How effective are long term mean and standard deviation as normalisation parameters for tonal fundamental frequency?,” Speech Communication **10**, 229–247.

Rostolland, D. (1982). “Acoustic features of shouted voice,” Acustica **50**, 118–125.

Scherer, K. R. (1974). “Voice quality analysis of American and German speakers,” Journal of Psycholinguistic Research **3**, 281–290.

Stoicheff, M.L. (1981) “Speaking fundamental frequency characteristics of nonsmoking female adults,” J. Speech Hear. Res. **24**, 437–441.

Takefuta, Y., Jancosek, E. G., and Brunt, M. (1972). “A statistical analysis of melody curves in the intonation of American English,” in *Proceedings of the 7th International Congress of Phonetic Sciences, Montreal* 1971, 1035–1039.

Titze, 1. R**.** (1989). “On the relation between subglottal pressure and fundamental frequency in phonation,” J. Acoust. Soc Am. **85**, 901–906.

Traunmiiller, H. (1988). “Paralinguistic variation and invariance in the characteristic frequencies of vowels,” Phonetica **45**, 1–29.

Traunmüller, H. (1994). “Conventional, biological, and environmental factors in speech communication: A modulation theory,” Phonetica **51**, 170–183.

Traunmüller, H., and Eriksson, A. (1995). “The perceptual evaluation of F_0-excursions in speech as evidenced in liveliness estimations,” J. Acoust. Soc Am. **97**, 1905–1915.

Williams, C. E., and Stevens, K. N. (1972). “Emotion and speech: some acoustical correlates,” J. Acoust. Soc Am **52**, 1238–1250.