

THE ESTIMATION OF INTRINSIC F_0 : A COMPARATIVE STUDY

D.J. Hirst, A. Di Cristo and Y. Nishinuma, Institut de Phonétique,
Université de Provence, France

A large number of studies have been devoted to the question of the intrinsic frequency (F_{0i}) of vowels in various different languages. These studies consistently indicate a strong inverse correlation between F_{0i} and the first formant of the vowel. The coefficient of determination (R^2) between F_1 and F_{0i} for the data given by Peterson and Barney (1952) is 0.85. Calculating the regression line from F_1 to F_{0i} consequently gives a reasonably close estimation of F_{0i} . This estimation can be considerably improved if we take into account the second formant (F_2), since we obtain an R^2 of 0.922. An even better correlation is found between F_{0i} on the one hand and F_1 , F_2 and $\overline{F_0}$ (the mean F_0 for each subject) on the other hand, ($R^2 = 0.976$) for the data from 11 different authors on 6 different languages. The estimation from the multiple linear regression on these data is very close to the original data ($r = 0.988$) and, although the correlation varies from author to author, in most cases the difference between the estimation and observed values rarely exceeds 2%.

A linear function $F_{0i} = a_0 + a_1 \overline{F_0} + a_2 F_1 + a_3 F_2$
where $a_0 = 20.166$, $a_1 = 0.975$, $a_2 = -0.034$, $a_3 = -0.002$ provides a very reliable estimation of the intrinsic frequency of vowels which can consequently be used both in prosodic analysis and in automatic speech synthesis and recognition.

Reference

Peterson, G.E. and H.L. Barney (1952): "Control methods used in a study of vowels", JASA 24, 175-184.