A REGRESSION ANALYSIS OF THE RELATION BETWEEN PHYSIOLOGICAL SIGNALS AND F0

Helmer Strik

University of Nijmegen, Dept. of Language and Speech
P.O. Box 9103, 6500 HD Nijmegen, The Netherlands
E-mail: strik@let.kun.nl

ABSTRACT

Measurements were obtained of several physiological mechanisms which are known to be important in the control of fundamental frequency (F0). The data were analysed by means of a multiple regression analysis in which F0 is the criterion and the physiological signals are the predictors. Separate analyses were carried out for statements and questions, and for falling and rising F0. The results reveal no considerable differences in the control of F0 for the various datasets.

1. INTRODUCTION

In the literature different views are expressed regarding the relation between F0 and the underlying physiological signals. The goal of the present research was to clarify this relation. Research on the relation between F0 and the physiological processes is very complicated, if only because F0 is dependent on a large number of physiological mechanisms [1]. Moreover, direct measurements of laryngeal physiology are by necessity invasive. Since these measurements are difficult to make, only a small amount of data is usually available.

To study the relation between F0 and the physiological signals, it seems advisable to use a quantitative analysis method. However, in most of the studies on this topic a kind of qualitative analysis is used. Two notable exceptions are [2] and [3]. Due to space limitations, it is not possible to go into the details of these two studies. Therefore, only the most important drawbacks of these studies are briefly presented here.

Both in [2] and [3] the total number of samples for which the quantitative analysis is done, is very small (i.e. 568 and 106, respectively). In these two studies analyses were also performed for subdivisions of the data. In these cases the number of data is even smaller. Another drawback of [2] is

that only correlation coefficients were calculated, and no regression equations. The reason why this is a drawback will be explained below. In [3] regression equations are presented, but in this study sustained phonation was used. It is not unlikely that the relations between F0 and the physiological signals in sustained phonation are different from the relations in running speech, as was already suggested in [3]; especially, because the F0 values found in [3] are very high (i.e. much higher than F0 values which are usually found in running speech).

In the current study measurements of physiological signals were made while subjects produced meaningful Dutch sentences. Our intention was to obtain a large amount of data, in order to have sufficient samples for the regression analysis.

2. MATERIAL AND METHOD

For two Dutch male subjects (LB and HB) recordings were made of the audio signal, electroglottogram, lung volume, subglottal pressure (P_{sb}), and the electromyographic activity of two laryngeal muscles: sternohyoid (SH) and vocalis (VOC). In addition to these signals, the activity of the cricothyroid (CT) muscle was also measured for subject LB, and oral pressure (P_{or}) for subject HB. The measurements were made while the subjects produced meaningful Dutch sentences with different intonation patterns. Each sentence was repeated 5 to 8 times. The signals of these repetitions were used to calculate average signals for every sentence. A more elaborate description of the experiments, and figures of the measured signals can be found in [1]. Here only those aspects are mentioned which are most relevant to the present article.

All signals were sampled at a 200 Hz rate, and were then smoothed. The muscle signals were shifted forward in time by their

Table 1. Results of SMRA for all data of subjects LB and HB. Shown are, from left to right, the regression coefficients C_i, the multiple correlation coefficient (MR), the number of datapoints (N), the identification of the regression equation (subject + number), and a brief description of the data.

C_0	C_1	C_2	C_3	C_4	MR	N	id.	description
67.9	2.9	0.061	0.378		0.886	2319	HB1	all data
67.1	3.2	0.063	0.376	-1.1	0.887	2319	HB2	all data
68.5	3.5	-0.21	0.444		0.836	2254	LB1	all data
70.7	2.7	-0.16	-0.01	0.51	0.896	2254	LB2	all data

mean response time (as described in [2]). Only the voiced frames of the utterances were used in a stepwise multiple regression analysis (SMRA). In the SMRA the dependent variable (the criterion) is F0, and the measured physiological signals are the independent variables (the predictors):

$$F_{0,est} = C_0 + C_1*P_{sb} + C_2*SH + C_3*VOC [+ C_4*X_4].$$

The fourth term in the regression equation (X_4) is only used once for each subject (see section 3.1). In that case X_4 is different for the two subjects, i.e. P_{or} for HB and CT for LB.

For different datasets correlation coefficients and regression equations were calculated. Furthermore, for each regression coefficient the standard error and the t-value were also computed. The t-values were used to check the statistical significance of the regression coefficients, while the standard errors were used to round off the regression coefficients to their last significant digit.

3. RESULTS

3.1. All data

First of all, regression equations were calculated for all data of both subjects. The results are given in Table 1, and the correlation coefficients in Table 2. A comparison of the regression equations HB1 and LB1 reveals that C_0 (the constant term), C_1 (the F0-P_{sb} ratio) and C_3 (the F0-VOC ratio) do not differ much between these subjects. However, their C_2's (the F0-SH ratio) are different. In most studies (see the references given in [1]) a negative relation between F0 and SH is found. The results of subject LB are in line with this general finding, but the results for HB are not.

Apart from P_{sb}, SH and VOC, other

physiological signals were measured for these subjects. For subject HB oral pressure (P_{or}) was also measured. The correlations of P_{or} with F0 are very small (for all 2319 voiced frames of HB the correlation is 0.011). Consequently, adding P_{or} to the regression equation does not have much influence. The resulting regression equation HB2 is almost equal to the regression equation HB1.

For subject LB the activity of the cricothyroid muscle (CT) was also measured. The correlations of CT with F0 are very high (for all 2254 voiced frames of subject LB it is 0.859). In fact, the correlation of CT with F0 is larger than any of the other correlations with F0 (see Table 2, row LB1). This is in accordance with what is usually found (see e.g. [2, 3]). The correlation between CT and VOC is 0.900 for all 2254 voiced frames. A high correlation between CT and VOC was also found by [2, 3]. Therefore, it seems that VOC acts in synergy with CT in the control of F0.

For subject LB the CT was added to the regression equation, and the result is equa-

Table 2. Correlations of F0 with Psb, SH and VOC for different subsets of the data.

Psb	SH	VOC	id.	description
0.333	0.351	0.872	HB1	all data
0.452	-0.404	0.760	LB1	all data
0.501	0.424	0.846	HB3	statements
-0.167	0.178	0.921	HB4	questions
0.594	-0.423	0.705	LB3	statements
0.167	-0.351	0.863	LB4	questions
0.191	0.404	0.834	HB5	falls
0.307	0.448	0.872	HB6	rises
0.601	-0.450	0.686	LB5	falls
0.320	-0.364	0.825	LB6	rises

Table 3. Results of SMRA for different subsets of the data. For explanation see Table 1.

C_0	C_1	C_2	C_3	MR	N	id.	description
67.8	3.5	0.079	0.334	0.871	1624	HB3	statements
64.7	3.0	0.04	0.434	0.925	695	HB4	questions
70.9	3.7	-0.26	0.42	0.808	1542	LB3	statements
58.7	3.7	-0.07	0.477	0.901	712	LB4	questions
78.1	1.7	0.07	0.325	0.848	586	HB5	falls
77.6	1.5	0.11	0.360	0.880	623	HB6	rises
61.8	4.8	-0.22	0.38	0.825	1101	LB5	falls
75.1	2.6	-0.21	0.46	0.873	484	LB6	rises

tion LB2 in Table 1. By comparing equation LB1 and LB2 it becomes clear that adding the CT has an enormous influence on the resulting regression equation. First of all, the multiple correlation increases substantially. Second, and more important, the magnitude of all regression coefficients changes. The reason why the changes are so considerable, is that the different variables are not orthogonal. This is certainly the case for CT and VOC. Consequently, a large part of the variance of F_0 that is explained by the VOC in equation LB1, will be explained by the CT in equation LB2. In equation LB2 C_3 (the F_0-VOC ratio) even becomes negative, while it is clear that F_0 and VOC are positively related.

This is an obvious disadvantage of regression equations. If the variables are not orthogonal, which is usually the case for physiological signals, the results of regression equations should be interpreted with caution.

Since P_{sb}, SH and VOC are the signals which were measured for both subjects, only these variables will be used in the rest of this article. Adding an extra variable (especially CT) does increase the amount of explained variance, but makes it impossible to compare the data between subjects. Because CT and VOC have similar effects on F_0, it is not so important which of the two variables is chosen.

After having calculated regression coefficients for all the data of both subjects, regression coefficients were computed for different subdivisions of the data: statements vs. questions, and falling vs. rising F_0. Similar subdivisions were made in [2], which makes it possible to compare the results of [2] with those of this study.

3.2. Statements and questions

In [2] the most striking differences between statements and questions were observed for the correlation of F_0 and P_{sb}. In statements it was positive, while in questions it was negative. The same effect can be observed for subject HB (see Table 2, compare rows HB3 and HB4). For subject LB the correlation of F_0 and P_{sb} is still positive for the questions, but it is much smaller than that for the statements (see Table 2, compare rows LB3 and LB4).

Although there are substantial differences between the correlations of statements and questions (also for the other variables, see Table 2), it can be observed in Table 3 (compare HB3 with HB4, and LB3 with LB4) that the differences between the regression coefficients are not so large. In other words, the regression equations reveal that the relations between F_0 and the physiological signals for statements and questions do not differ much. This is an example of an advantage of regression analysis compared to simple correlation analysis. Even if the relations among the variables are almost the same (i.e. the regression coefficients are almost the same), the correlation coefficients can have very different values depending on the kind of data used (e.g. statements vs. questions).

3.3. Falling and rising F_0

In the section above, the data were divided into statements and questions. In this section the data will be subdivided in terms of falling and rising F_0. Samples with a negative derivative are classified as falls, and samples with a positive derivative as rises. The same method was also used in [2], which makes it possible to compare the results.

For subject HB the correlation between F_0 and P_{sb} is different for falls and rises (see Table 2, compare rows HB5 and HB6), whereas the other correlations and the regression coefficients are very similar (see Table 3, compare rows HB5 and HB6). For subject LB larger differences between falling and rising F_0 can be observed, both for the correlations and the regression coefficients (see Tables 2 and 3, compare rows LB5 and LB6). In [2] the largest difference between rises and falls was also found for the correlation between F_0 and P_{sb}, as in our data. For the other correlations no substantial differences were observed in [2] (except for a difference for the correlation of F_0 and the lateral crico-arytenoid muscle).

In short, differences between falls and rises are observed in the correlations of F_0 and P_{sb} for all subjects, and in the regression coefficients of subject LB. In the latter case those differences are particularly evident for C_1 (the F_0-P_{sb} ratio).

4. DISCUSSION

In this article I have presented the results of a quantitative analysis of the relation between F_0 and some physiological mechanisms that are known to be important in the control of F_0. First of all, it is important to note that (apart for the coefficients for P_{or}) all correlation and regression coefficients are highly significant, reflecting the consistent relations among the variables. This was also found in [2] and [3].

The analysis results for all data show that the effect of the SH on F_0 was different for the two subjects, but for the other variables no major differences were found. The variables showing the highest correlations with F_0 were CT and VOC. The correlations of F_0 with P_{sb} and SH were always smaller.

Substantial differences were found between the correlation coefficients calculated for statements and questions, but the differences in the regression coefficients were not very large. Comparing the analysis results for falls and rises revealed that there were differences in the correlations of F_0 and P_{sb} for all subjects, as well as in the regression coefficients of subject LB (especially for C_1, the F_0-P_{sb} ratio). Whether these differences should be interpreted as large, remains questionable. More research

is needed to give a definite answer to this question. For the time being, my interpretation of the results is that the relation between F_0 and the physiological signals in statements and questions, and in falls and rises is not very different.

The advantages of the present study, compared to [2] and [3] are, that the number of samples is much larger, that the measurements were obtained for running speech, and that besides correlation coefficients also regression equations were calculated. As mentioned above, regression coefficients are sometimes preferable to correlation coefficients. The reason is that for some subsets of the data the correlations are very different, while the regression coefficients (and therefore probably the underlying relations) are very similar. However, when the variables are not orthogonal, one should also be careful in interpreting the results of regression equations.

In the regression analyses carried out in this study, the physiological signals were used as independent variables (the predictors). Given that no explicit model was used, the implicit assumption made by using this analysis method is that the relation between F_0 and the physiological signals is linear. However, it is almost certain that this relation is not linear. For a more realistic modelling of the relation between F_0 and the physiological processes a production model is needed in which not only the vocal tract but also the voice source is modelled in a physiologically meaningful way. At the moment, a model of this kind does not exist. More research is needed to develop and test such models.

REFERENCES

[1] Strik, H. (1994) *Physiological control and behaviour of the voice source in the production of prosody*. Ph.D. thesis, University of Nijmegen.

[2] Atkinson, J.E. (1978) Correlation analysis of the physiological features controlling fundamental voice frequency. *Journal of the Acoustical Society of America*, 63, pp. 211-222.

[3] Shipp, T., Doherty, E.T. & Morrissey, P. (1979) Predicting vocal frequency from selected physiologic measures. *Journal of the Acoustical Society of America*, 66, pp. 678-684.