(ARIPUC 15, 1981 p. 103-146)

THE EFFECT OF INCREASED SPEAKING RATE AND (INTENDED) LOUDNESS ON THE GLOTTAL BEHAVIOUR IN STOP CONSONANT PRODUCTION AS EXEMPLIFIED BY DANISH P.

PREBEN ANDERSEN

The present combined fiberoptic-glottographic experiment has the purpose of studying the glottal behaviour in Danish stop consonant production at increased speaking rate and intended loudness, respectively. At increased speaking rate the degree and duration of the glottal abduction is found to decrease proportionally, i.e. the rate of the glottal abduction is constant, irrespective of speaking rate. Different individual strategies seem to be applied in the temporal organization of the glottal adduction. At increased intended loudness level no effect on the temporal organization of the gesture is observed, but the rate of glottal abduction is found to be faster. A model is proposed for the glottal behaviour under each of the two speech conditions. A very limited EMG-material (1 subject) shows changes in INT-activity under both conditions, whereas changes in PCA-activity occur only in connection with increased intended loudness level. The present findings concerning the speaking rate condition are compared to results from research on supraglottal articulatory organization at varying speaking rates.

I. INTRODUCTION

The present study originated from observations made in a fiberoptic-glottographic pilot study. Glottograms of repetitive productions of stops in conversational speech displayed a considerable variation in maximum amplitude, and supplementary acoustic curves displayed variation also in the duration as well as in the intensity of the open interval of the stops.

The present investigation can be considered a preliminary contribution to the determination of a potential systematic relationship between degree of maximum glottal abduction and (1) intensity and (2) speaking rate (as reflected by the duration) of stops in conversational speech. Its specific purpose is to clarify the effect of deliberate variations of intensity level and speaking rate on the glottal behaviour of Danish stops. The application of the results, presented here, to the normal, conversational speech condition will be the subject of a later paper.

This paper has another, perhaps more important purpose, namely to contribute to the debate on timing in speech production. It is commonly known that allophonic variation due to the suprasegmental features of stress and speaking rate among other things includes durational differences. The majority of research in this area involves an attitude to a model, proposed by Björn Lindblom in 1963 (Lindblom 1963). This model assumes that the allophonic variation in destressed vowels or vowels produced with faster speaking rate is merely a consequence of the shorter duration and can be accounted for as a pure timing phenomenon. For vowels subject to variation in stress and speaking rate, Lindblom found a high correlation between the duration of the vowel and an expression of its degree of spectral reduction, i.e. the shorter the vowel, the closer its formant pattern to that of schwa. Lindblom concludes that this neutralization is a consequence of articulatory and acoustic undershoot, caused by a temporal overlap of the muscle commands.

The essential question raised in connection with Lindblom's model has been, whether a model of destressed speech and speaking rate control can be based on timing changes alone. Later experiments (Harris 1971, 1978, Kent and Netsell 1971, Gay 1978) have given evidence that production of vowels under varying stress conditions seems to involve a change in the intensity of the muscle commands as well and, at the peripheral level, a change in the rate of articulator displacement. At the muscular level, Sussman and MacNeilage (1978) and Sussman (1979) have observed an increase in the number of active motor units during the production of stressed syllables and, at the individual motor unit level, a higher instantaneous discharge rate. A conclusion from the results of recent research has been (Harris 1978), that control of vowels under different stress conditions seems to be explicable in terms of an "extra energy model" (Öhman 1967) rather than of Lindblom's "undershoot model".

For vowels of different duration caused by varying speaking rate, the mechanism of the underlying control seems less evident. Gay et al. (1974) observe for vowels produced at faster speaking rate an undershoot in tongue displacement, a decrease in tongue EMG-activity, but no concurrent decrease in articulator velocity. They conclude that although an articulatory undershoot is stated, one cannot fit in the findings with Lindblom's model, as this presupposes the undershoot to be a mere consequence of temporal overlap of muscle commands with no changes in driving force. As regards velocity of articulator displacement, Kuehn (quoted from Gay et al. 1974) has shown that different speakers use different strategies to increase speaking rate, e.g. by trade-offs in displacement vs. velocity. Harris (1978) found longer and greater EMG-activity for vowels in stressed position (see above) but no consistent effects on EMG-activity from speaking rate. In a spectrographic study Gay (1978) found variations of vowel-target frequency as a function of stress, but not as a function of rate. He concluded that control of stress and speaking rate implies different mechanisms.

As Lindblom's model accounts for vowels only, an essential question is whether the complex of problems connected with the model can be applied to consonant production as well, i.e. does consonant production, when influenced by changes in stress and speaking rate, imply articulatory reorganization comparable with that observed for vowels? The effect of speaking rate on labial consonant production has been studied by Gay et al. (1973, 1974). Besides differences in the timing of labial muscular events they found an increase in labial muscle activity and in rate of labial movement by increased speaking rate. Accordingly, their conclusion was that a general model of speaking rate control, based solely on timing changes, is too simple. The concept of undershoot was reconsidered, as vowel production involves articulator movements towards spatial targets, whereas consonant production most often implies movements towards occlusive or constrictive targets. Consequently, consonant production is supposed not to allow undershoot, and the authors find it reasonable to assume that consonant production, when constrained by increased speaking rate, implies faster movement of the articulators towards the target.

As for consonant production under different stress conditions, Harris et al. (1968) and data from other investigations (Lubker and Parris 1970) suggest no consistency as concerns differences in driving force, although consonants in pre- and post-stressed position have different durations.

From the above it is evident that research on mechanisms of stress and rate changes has been primarily concerned with the control of the supraglottal articulation. The second purpose of the present paper is to contribute further to the complex of problems connected with articulatory organization under different stress and tempo conditions by shedding light also on the glottal articulation. The study is concerned specifically with the glottal behaviour during stop consonant production under different speech tempo conditions. The essential question raised in connection with the experiment is, if an increase of speaking rate also implies a reorganization of the glottal gesture in terms of faster rate of ab- and adduction and hence, underlyingly, a change in the timing and intensity of the commands to the laryngeal muscles? This is part of a more general question: Are supraglottal and glottal articulations during consonant production under different speaking rate conditions controlled by one and the same underlying mechanism? An alternative would be that supraglottal and glottal articulations in this respect are independently controlled and work as two different systems. The latter possibility seems plausible if the concept of target undershoot is taken into consideration. Whereas the supraglottal articulation implies articulator movement towards an occlusive target, the glottal articulation might be perceived as a movement towards a far less constrained spatial target, the nature of which would be expected to allow articulatory undershoot in contradistinction to that of the supraglottal target.

Experimentally the problems are approached here through extensive fiberoptic-glottographic recordings in order to account for the dynamic properties of the glottal gesture, supplemented by a very limited EMG-material, supposed to yield some information about the mechanisms causing the dynamic variation.

II. METHOD

A. SUBJECTS AND LINGUISTIC MATERIAL

The speakers were two adults, a female (subject HU) and a male (subject JR).

The linguistic material for the investigation was limited to the Danish stop *p* in covered initial position. The test sound was embedded in the phrase *de vil sige pige* [d̥i ve si: ˈb̥ʰi:] 'they will say (the word): girl'. The articulatory course of the phrase was supposed to minimize alterations in the position of the fiberoptic cable.

The material to be recorded consisted of repetitions of the test phrase under the following two speech conditions:
a) in speech where the subject, repetition by repetition, in subjectively determined steps, was told to gradually speak more loudly, starting the series with the phrase being produced with weak "loudness". (The change in the subject's performance as a consequence of this instruction is here referred to as a change in "loudness".) Each subject spoke 9 such series (in the following referred to as '"loudness" series').
b) in speech where the subject, following the same principle as described above in a), produced series of repetitions in which the speaking rate of the phrase was gradually increased. Subject HU spoke 9 such series, subject JR 8 series (in the following referred to as 'rate series').

B. THE INSTRUMENTAL SET-UP

The fiberscope (Olympos VF, type 4A) and the photo-electric glottograph (a slightly reconstructed version of the one described by Frøkjær-Jensen 1967) are the central components of the instrumental set-up shown in figure 1. The light guide of

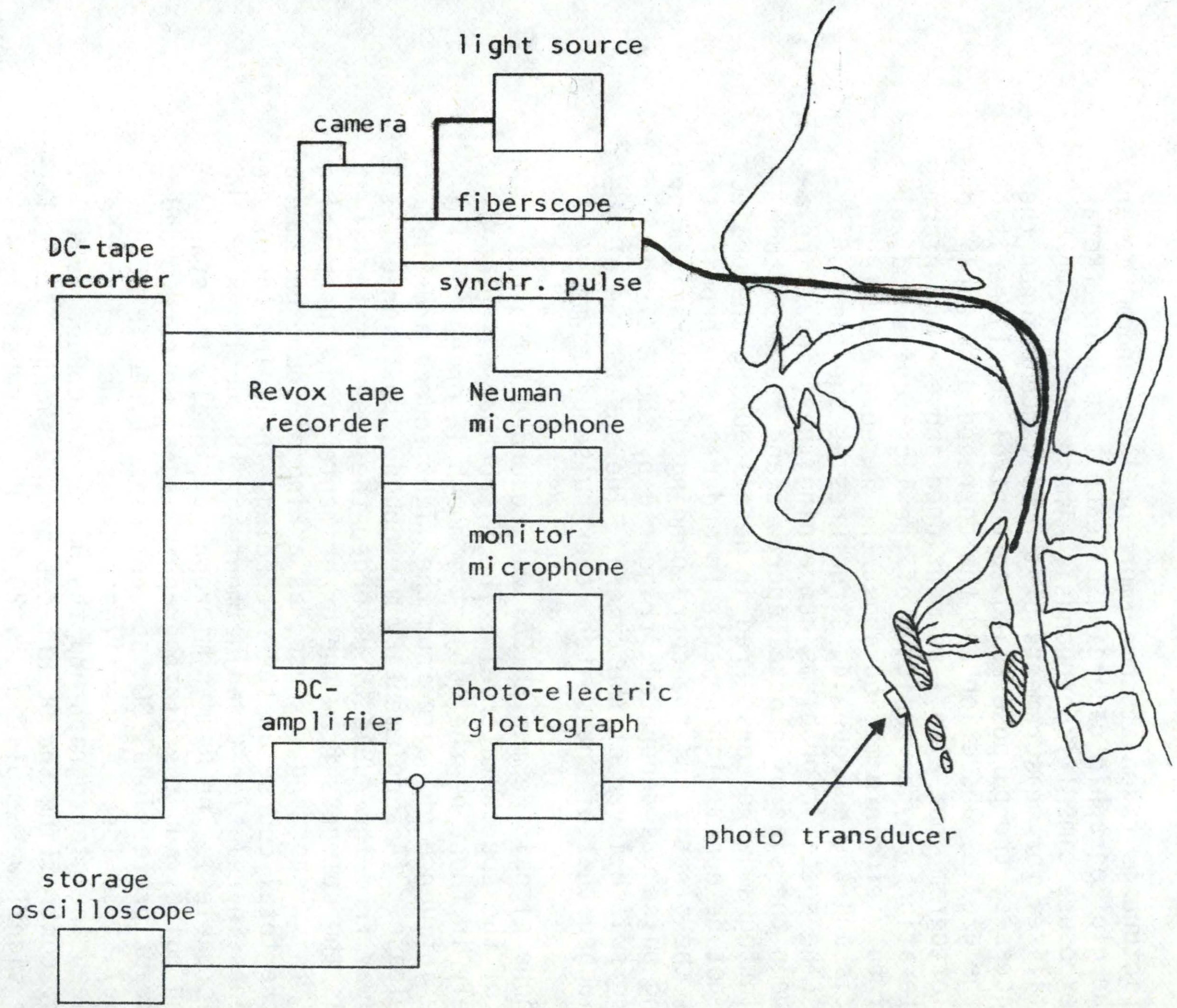


Figure 1

The instrumental set-up.

the fiberscope, placed in the subject's pharynx, serves as the light source for the photo-electric glottograph. This technique makes it easier to control the position of the light source by means of the fiberoptic pictures transmitted through the image guide. A photo transducer, placed on the frontal part of the subject's neck in a position approximately between the thyroid and cricoid cartilages, turns it into an electric voltage variation which, after amplification, can be recorded graphically.

A still-picture is taken at some (more or less random) moment during the glottographic recording of each test sound (Kodak TX 135, exposure time 1/125 seconds). These pictures were mainly utilized for measurements of the distance between the vocal processes, the purpose being to control the linearity between the glottal abduction and the corresponding amplitude of the glottogram, but also - as mentioned above - to inform about alterations in the position of the light guide and, altogether, to yield information about the immediate state of the glottis and the surrounding structures at the time of exposure. (The specific problems concerning the (linear) relation between the size of glottal aperture and the amplitude of the glottogram were considered in detail during the project, but will not be dealt with here.) In order to synchronize the stills of the glottis and the corresponding glottograms, a synchronizing pulse was generated, triggered by the M-X synchronizing switch lever of the camera. The onset time of the shutter lagged that of the switch lever pulse by 15 ms.

A microphone signal (Neuman KH 56) was recorded for acoustic analysis of the speech material. The microphone was placed immediately in front of the subject's mouth in order to eliminate the influence of noise from the light source and the camera release. (A supplementary microphone, positioned farther away from the subject, recorded the speech material as well as the comments from the experimenter and others being present. This recording was used as a kind of sound documentation of the total course of the recording, including the pauses and comments in relation to adjustments of the position of the fiberoptic cable.) The Neuman microphone signal, the synchronizing pulse, and the glottographic signal were recorded on an 8-track professional DC tape recorder (Lyrec TR 86, Agfa PE 36) at 30 ips. An oscilloscope was connected to the line between the glottograph and the DC tape recorder as a device for controlling the DC off-set of the signal. The glottographic signal was amplified by 20 dB (PT Universal Amplifier) for both subjects.

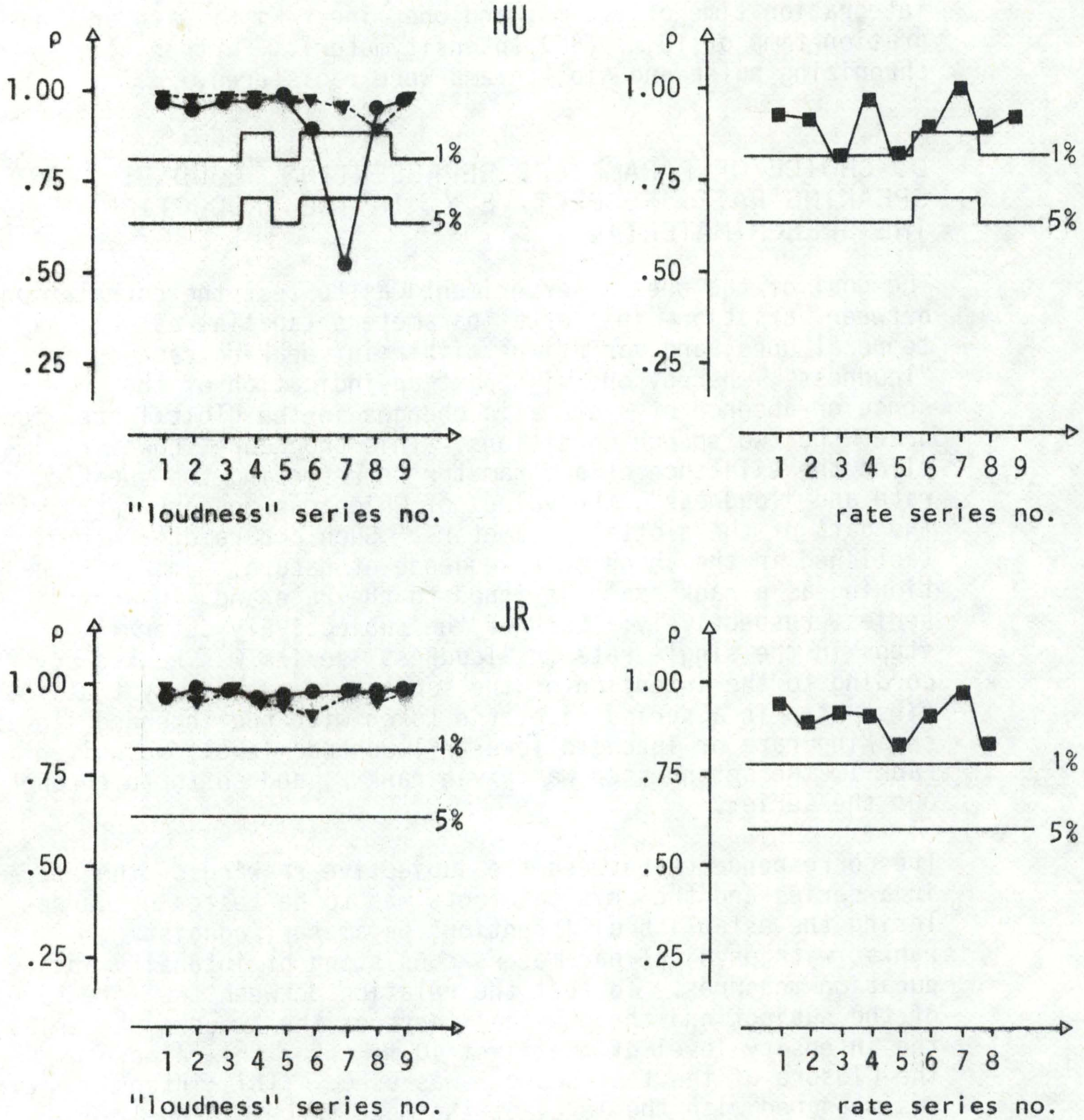
C. GRAPHICAL REGISTRATION OF THE RECORDED MATERIAL

The speech material was registered graphically on the mingograph (paper speed: 100 mm/sec.). From the Neuman microphone recording the following curves were extracted for further analysis: Duplex oscillogram (BFJ Transpitchmeter); 2 intensity curves, one logarithmic HP-filtered at 500 cps with an integration time of 2.5 ms, and one linear Hi-Fi with an integration time of 10 ms (BFJ Intensitymeter). Further, the synchronizing pulse and glottograms were registered.

D. CHOICE OF PARAMETERS REPRESENTING "LOUDNESS" AND SPEAKING RATE, RESPECTIVELY, IN THE PRODUCTION OF THE SPEECH MATERIAL

The goal of the present experiment was to test the correlation between variations in glottal parameters (spatial as well as temporal ones) and variations either in speaking rate or in "loudness", whereby one might get an indication of the presence or absence of systematic changes in the glottal behaviour under the two speech conditions. This procedure, however, implies the existence of a parameter representing the speaking rate and "loudness", the values of which can be correlated with the data of the glottal parameters. Such a parameter was established in the shape of a sequence of natural numbers functioning as a rank scale assigned to the rate and "loudness" series, respectively. Each of the subjectively determined steps in the single rate or "loudness" series was ranked according to the intention of the subject in such a way that the first step in a series (i.e. the token with the intended slowest speaking rate or intended lowest "loudness" level) was given rank 1, the second step was given rank 2, and so forth throughout the series.

The correspondence between the subjective ranking of the steps in a series and the physical facts was to be tested by correlating the established 'intention' parameter, consisting of ranks, with physical parameters consisting of intensity and duration measures. To test the relation between the intention of the subject and the physical facts of the "loudness" series, the intensity level of the first 40 ms after the release of the closure of the test sound *p* was used. (This intensity level was measured with the level of the most forceful vowel in the material as reference.) To test the corresponding relations as regards the rate series, the acoustic duration of the stop (closure + open interval) was measured. The intensity and duration data were converted to ranks, and for each subject the Spearman's coefficient of rank correlation between the 'intention' parameter and each of the physical parameters was computed. The results of the computation is shown in figure 2. As regards the "loudness" series, the rank correlation coefficient for the 'intention' parameter and the average intensity level of the vowel *i* following *p* in the the test phrase is included in the figure in order to give an impression of the intensity variations not only in the test consonant. It is evident that there is - within each series - a highly significant, positive correlation between the intentions of the subjects and the physical facts. As the aim of the present experiment was to state, for the test sound, the relations between its "loudness" and rate variations, respectively, and its glottal behaviour, not only in terms of correlation coefficients but, as far as possible, also in terms of physical,



● = rank correlation between 'intention' and intensity of *p*
▼ = rank correlation between 'intention' and intensity of *i*
■ = rank correlation between 'intention' and duration of *p*

Figure 2

Spearman's coefficient of rank correlation (ρ) for the 'intention' parameter and (a) the intensity of *p* (intensity level of the first 40 ms. after the release of the closure) ,(b) the intensity level of the *i* following *p*,(c) the duration of *p* (closure + open interval). The first two coefficients are calculated for each of the "loudness" series, and the third one for each of the rate series. Two levels of significance are included.

numerical relations, the two physical parameters were chosen to represent the rate of performance of the sound and thereby the variations in respectively the speaking tempo and the variations in "loudness".

1. AVERAGE "LOUDNESS" AND RATE SERIES

Figures 3a and 3b illustrate graphically the measured intensity of the explosion and aspiration of *p* through each of the 9 recorded "loudness" series and the measured segment duration of *p* through each of the 9 (subject HU) or 8 (subject JR) recorded rate series. Though there is some dispersion of the values on each step which causes a certain amount of overlapping between consecutive steps, an average increase in intensity and decrease in duration through the series is obvious for both subjects. One might suspect that the dispersion of the values on each single step is caused by systematic differences in the course of intensity increase or duration decrease between the series, for instance in the sense that one series was dislocated one step in relation to another. However, no such clear parallelism in the course of any two series is observed, and one must conclude that the dispersion of data is random for each step. This fact makes it possible to pool the series and, thus in the following, to operate with only one average "loudness" and average rate series. (The average value of each step is indicated on the figure by a small horizontal line.)

In the rate series, for subject JR each one spanning 9 steps and for subject HU only 8 steps, it can be observed that through the series subject JR shortened the segment duration by 100 ms, on the average, subject HU by only 73 ms on the average. The total reduction, however, differed for the two subjects with a value just about equal to the range between the first and second step in subject JR's average series, and as the second step in the average series of subject JR had a value corresponding to that of the first step in the average series of subject HU, one might tentatively conclude that subject JR started his rate series one step slower than did subject HU. In any case, for the sake of comparison, the first step in the material of subject JR was left out in the further analysis.

The values of the steps in the average series were used as the parameters representing "loudness" and rate, respectively, in the further analysis. In the following, the two parameters constitute the standard of reference in the testing of the glottal and partly the supraglottal behaviour of *p* under the influence of varying "loudness" and speaking rate. The parameters to be compared (i.e. correlated) with the "loudness" and tempo parameter will be described in the following section. In accordance with the "loudness" and tempo parameter, each of these spatial and temporal parameters will also be represented by only one value per step, namely the average value of the step in question for all the series recorded.

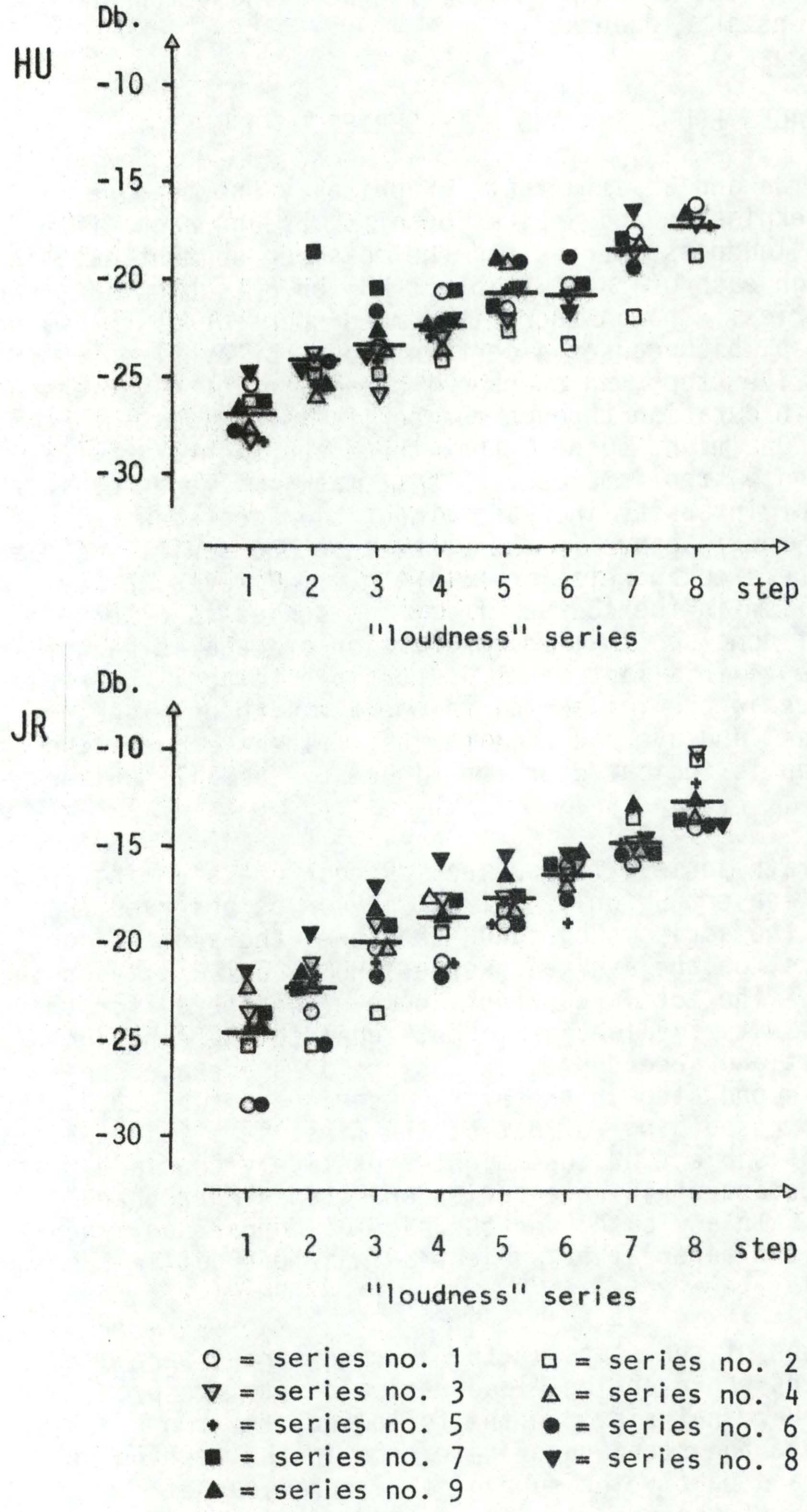


Figure 3a

The measured values of the intensity level of *p* (intensity level of the first 40 ms. after the closure release) throughout the "loudness" series.

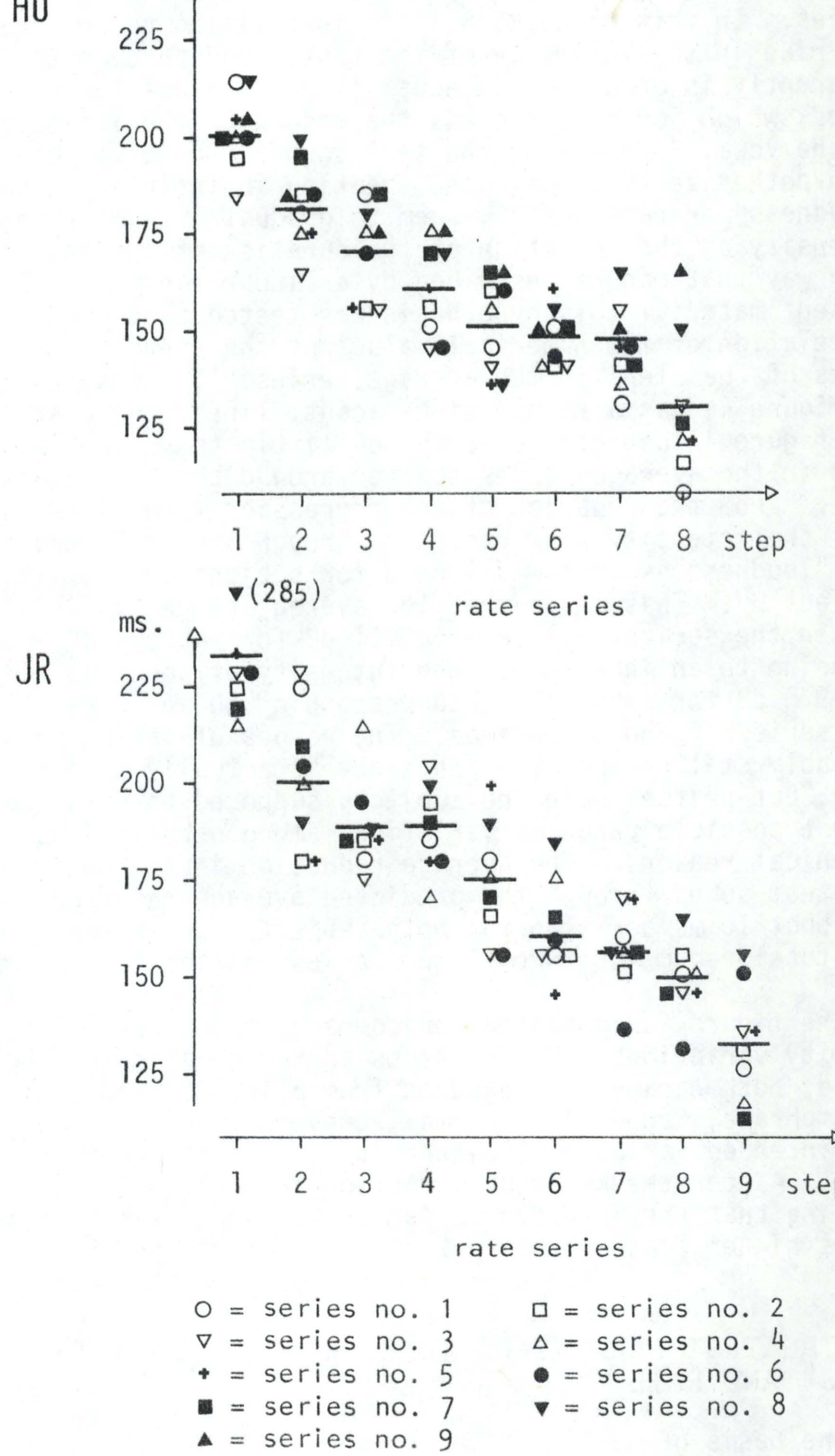


Figure 3b

The measured values of the duration of *p* (duration of closure + open interval) throughout the rate series.

2. PREDICTED AVERAGE "LOUDNESS" AND RATE SERIES

For neither subjects is there any obvious tendency towards a systematic difference in the average intervals between the single steps, not through the "loudness" nor through the rate series. In this connection it is interesting to observe that the rise in the intensity of the test sound through the series apparently is graduated in equal linear and not logarithmic steps, which for instance was the case with the intensity rise of the vowel *i* following the test sound. One might consequently hypothesize that subjects, speaking an infinite number of "loudness" or rate series, seem to graduate the rise in the intensity or the reduction of the duration of the test sound in a way that can be described by a linear function. In the present material this hypothesis was tested by product-moment correlation of the numerical values of the steps with the ranks of the steps in the average series. The result, shown in figure 4, was a highly significant, linear correlation. The figure illustrates clearly how little the values of each step in the average series scatter around the Y'-regression line. From the equation of the regression line we can derive that the intensity rise per step through the predicted average "loudness" series was 1.6 dB for subject JR and 1.3 dB for subject HU. This means that the average range of variation within the series will be about 11 dB for subject JR, corresponding to an increase in the intensity of about 13 times, and 9.5 dB for subject HU, corresponding to an increase through the series of about 10 times. The ranges of variation are probably smaller than the range one is actually able to perform, but neither were the subjects supposed to span their widest possible range of variation, among other things for technical reasons. The average reduction in the duration of the test sound through the predicted average rate series will be about 10 ms per step for both subjects, which means that the total reduction through the series was about 70 ms or 50%.

On the figure is indicated for comparison the range of intensity variations and variations in the duration of the test sound, both parameters measured from *p* in 150 samples of the test phrase, repeated in normal, conversational speech without intended variation, neither in "loudness" nor in speaking rate. As for the range of variation in duration, it seems surprising that the intended variation mainly consists in speaking rates higher (rather than lower) than normal speech.

E. THE TEST PARAMETERS TO BE CORRELATED WITH "LOUDNESS" AND RATE

On the basis of delimitations on the acoustic curves and the glottograms, 5 temporal parameters, all measured in milliseconds, and one spatial parameter were extracted and measured for correlation with the rate parameter and the "loudness" parameter. Of the extracted parameters, 2 were purely acoustic, and 4 purely physiological (glottographic). The extracted parameters were the following:

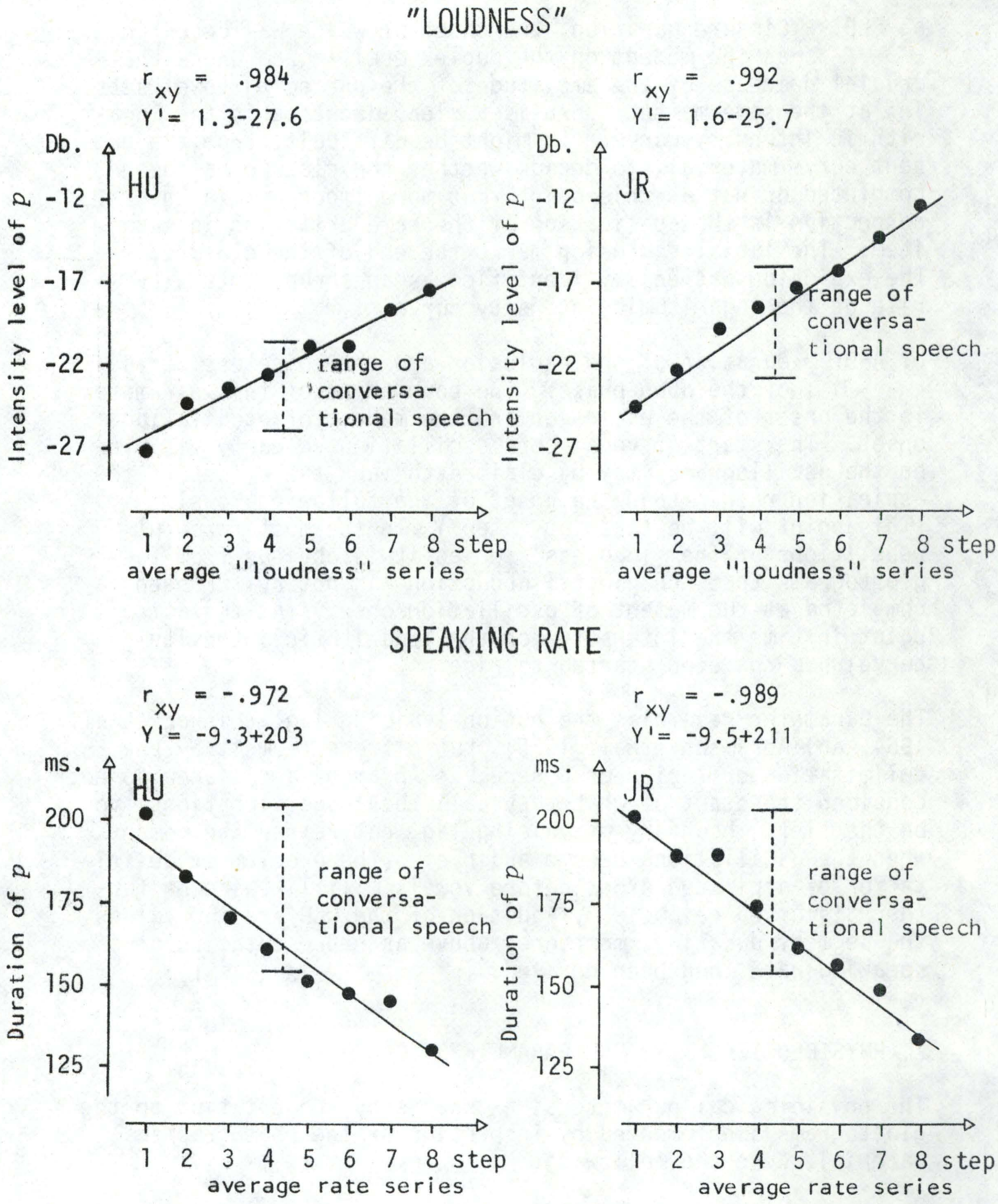


Figure 4

Intensity level of *p* on each step in the average "loudness" series as a function of the rank of the step (upper half) and duration of *p* on each step in the average rate series as a function of the rank of the step (lower half).

1. ACOUSTIC TEMPORAL PARAMETERS:

a. CLD: Closure duration, the onset of which was determined as the moment on the duplex oscillogram where the crucial decrease of the amplitude of the preceding vowel sets in; at the same moment there is a clear decrease in the logarithmic intensity curve. It might be difficult, from the present curve material, to decide whether the closure has been completed or not at this point, but more important in this connection is the application of the same criterion in each item. The labial explosion marks the end of the closure. The explosion was easily identified as an abrupt intensity rise on the logarithmic intensity curve.

b. ASD: Duration of the explosion and the voiceless period of the open phase. The boundaries of this parameter is the onset of the explosion and the moment of oscillation onset. The start of vocal cord oscillation, clearly visible on the oscillograms, may be claimed to mark the offset of the aspiration rather than the onset of the following vowel. (This point will be taken up later.) In the most forceful repetitions of the "loudness" series it is obvious from the glottograms that the glottal abduction has not by far been completed at the moment of oscillation onset. At the same point in time the intensity (on the logarithmic intensity curve) has not even started to rise.

The parameter resembles the notion 'voicing lag' (Lisker et al. 1964, and Abramson et al. 1965), but differs from it by the delimitation criterion used here, as Abramson and Lisker do not consider the start of what they call the 'edge vibrations' to be the 'left' boundary of voicing lag, but rather the moment when the oscillations become audible. (The problem of delimitation of aspirated stops before vowels will likewise be further discussed below.) By addition of the CLD and ASD values the segment duration, mentioned above as representative of speaking rate, had been derived.

2. PHYSIOLOGICAL TEMPORAL PARAMETERS

The physiological parameters, extracted by segmentation on the glottograms supplemented by inspection of the photographic material, were the following:

a. ABD: Duration of the glottal abduction of the test sound. The moment of completion of the abduction is defined as the moment of maximum amplitude on the glottogram, which was rather unproblematic to identify. As for the definition of the start of the abduction, this moment was taken to be the moment of initiation of the apparent movement of the arytenoid cartilages (deduced from fiberoptic pictures and glottograms) and the moment of initiation of the glottal gesture. A preceding check of the stability of the minimum level of the glottograms showed that this level to a high degree seemed to reflect no movement of the arytenoid cartilages and, translucency and incomplete closure excepted, approximately full adduction. On the glottograms the initiation of the glottal abduction was defined as the moment when the minimum level of the glottogram started to rise. An inspection of those stills that happened to have been taken about the moment of change in the minimum level of the glottogram, confirmed that the moment in question was a reasonable criterion for the determination of the start of the glottal abduction. Around this point in time the stills displayed a transition from a state of glottal vibration to the initiation of an active glottal abduction, starting posteriorly.

It shall be mentioned that, for the glottograms of the fastest repetitions in the rate series, it was not always unproblematic to identify the start of the glottal gesture as the minimum level of the glottogram which, because of the fast speaking rate, never reached a steady state level but displayed a kind of undershoot effect. In such cases the initiation of the glottal gesture of the test sound was defined as the moment when the minimum level of the glottogram of the preceding vowel was at its lowest.

b. ADD: Duration of the glottal adduction of the test sound. The onset is defined as the moment of maximum glottogram amplitude, and the completion as the moment when the falling flank of the glottogram has reached the minimum level. The determination of criteria for identifying the completion of the adduction were, as mentioned above, concerning the moment of start of the glottal abduction, also here supplemented by an inspection of the stills, taken "at random" close to the point in time of the moment in question.

c. OADD: Duration of the oscillating phase of the glottal adduction. The onset of the oscillations is clearly visible on the decreasing flank of the glottogram as well as on the duplex oscillogram, and it coincides with the offset of the acoustic parameter ASD, mentioned above.

3. THE SPATIAL PARAMETER

Besides the temporal parameters the maximum amplitude of the glottogram of the test sound was included as a spatial parameter. The maximum amplitude of the glottogram (MAG) was measured with the minimum level of the glottogram as a reference. It shall be mentioned that the results, presented in the following, are based on MAG-values, which are not all identical with the values actually measured on the glottograms. A common problem during fiberoptic-glottographic recordings are disturbances in the transmission line between the light source (i.e. the fiberoptic cable) and the photo transducer. The disturbances are caused by the subjects occasionally having to cough or swallow, or by the speech condition itself (e.g. during forceful speech the tip of the fiberoptic cable often changes position in relation to the glottis), etc. These sources of error will cause variation in the glottogram amplitudes that are not reflections of variation in the size of the glottal aperture. Although a highly positive correlation had been found between the size of glottal aperture, measured on the stills, and the glottogram amplitude, measured at the moment when the still was taken, sources of error might nevertheless have affected the recording, as the errors are not necessarily reflected in the degree of correlation between size of glottal aperture and glottogram amplitude but often in the slope of the regression line, describing the relationship between the two parameters. For instance, an approximation of the light source to the glottis (this can be ascertained from the stills) will intensify the exposure of the glottal opening and the photo transducer and thereby cause an increase in the slope of the regression line describing the relationship between glottal aperture and glottogram amplitude.

A normalization of the glottogram amplitude data was attempted on the basis of a closer inspection of the changes in the slope of the regression line throughout the recording procedure, compared with changes in the experimental conditions, established from the still picture material and from listening to a tape recording of the total recording process, including comments from experimenter, subjects, and others being present. Sections of the recordings, where the slope of the computed regression line was observed to be deviant from that of an arbitrarily chosen 'normal' section of the recording, were taken as a starting point in the normalization procedure. To the extent that the change to a deviant slope of the regression line could be plausibly accounted for by the still picture material and the control tape, the deviant amplitude data were normalized, i.e. their position in relation to the regression line about which they would have scattered if the sources of error had not affected the recording, was calculated.

The procedure of normalizing glottogram amplitude data, as applied here, should be considered only as a preliminary approach to a problem of methodology involving quite a number of uncontrolled factors. Naturally, the normalization procedure cannot account for all sources of error, and hypercorrections have probably been made (see below). However, the general impression is that this attempt to eliminate the influence from sources of error showed plausible results in most respects, based on reasonable assumptions about the nature of the sources of error and their expected influence, supplemented by physical evidence. On the whole, it seems warranted to assume that the normalized amplitude values reflect the physical facts to a higher degree than the originally measured ones, and thus they offer a better basis for pooling glottogram amplitude values of the repetitive recordings.

In the following, the average maximum amplitude through the "loudness" and rate series is represented by the median, as the values of each step displayed no underlying normal distribution. The 6 parameters are illustrated by examples in figure 5.

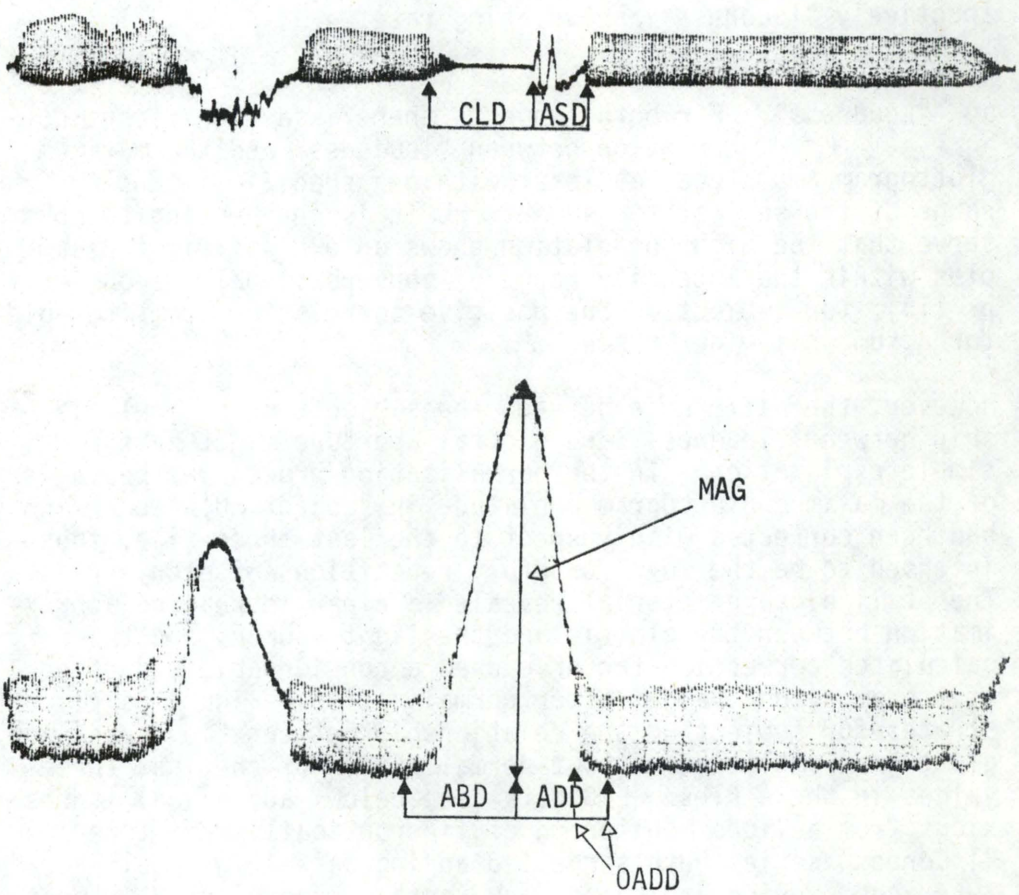


Figure 5

The 6 test parameters to be correlated with "loudness" and rate. The 2 acoustic parameters are illustrated in the upper curve (duplex oscillogram), and the 4 fysiological parameters are illustrated in the lower (photo-electric glottogram).

III. RESULTS

A. THE EFFECT OF "LOUDNESS" AND SPEAKING RATE ON THE SPATIAL REALIZATION OF THE GLOTTAL GESTURE

Figure 6 illustrates graphically for each subject the correlation between the maximum amplitude of the glottogram and respectively "loudness" and speaking rate.

a. "Loudness" For both subjects there is a significant correlation between "loudness" and the maximum glottogram amplitude, at least within a specific intensity range of the series; for subject HU it is interesting to observe that the high correlation shows up exclusively for samples within the intensity range of conversational speech (cf. p. 115); for subject JR the positive correlation seems to hold for a somewhat wider range.

However, the difference between the subjects in the relationship between "loudness" and glottal aperture might have a simple explanation. In the normalization procedure the values of the maximum glottogram amplitude in subject HU's recording had been corrected with respect to the last three (i.e. those intended to be the most forceful) repetitions of each series. The still picture material revealed a clear increasing approximation between the glottis and the light source, and the calculated correction factor caused a considerable reduction of the amplitude values after normalization. Figure 6a illustrates for subject HU the relationship between "loudness" and glottogram amplitude without normalization of the last three values in the series. EMG data (see below) and visual impressions from a Video monitoring of fiberoptically registered "loudness" series both suggested an increase in glottal aperture accompanying increased "loudness". Accordingly, it must be assumed that figure 6a in fact reflects the actual relationship between size of glottal aperture and "loudness" better than the normalized data. The increasing approximation between the glottis and the light source, ascertained from the still picture material, and the observed radical increase in the slope of the regression line taken into consideration, it seems plausible that the calculated small amplitude values in question are due to the presence of an additional source of error rather than the use of too high a correction factor.

It is evident from the figure that the relationship between "loudness" and glottogram amplitude of the non-corrected values for subject HU is much more like that of subject JR. For this subject no correction was made for differences in the distance between the glottis and the light source, caused by the "loudness" condition itself, although a normalization of the value of the last step in the series was considered.

Although the normalization procedure used here seems to warrant a more adequate description of the physical facts (see Note 1),

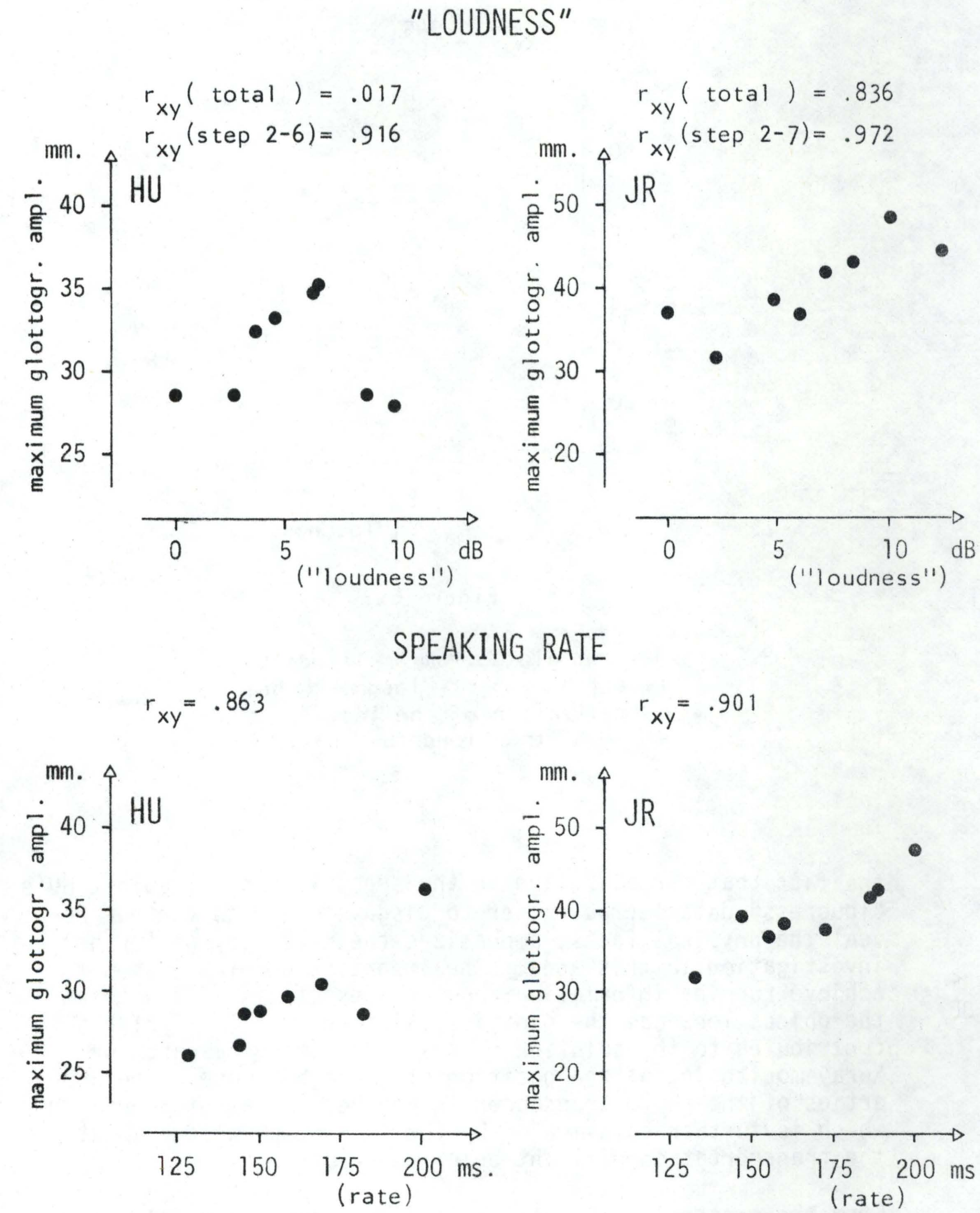


Figure 6

Maximum glottogram amplitude of *p* as a function of "loudness" (intensity of *p*) and speaking rate (duration of *p*).

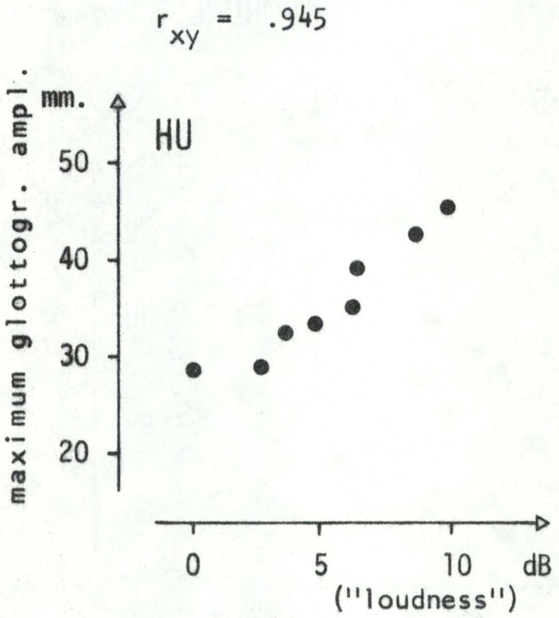


Figure 6a

Maximum glottogram amplitude of *p* as a function of "loudness" before normalization of the last three steps of the "loudness" series.

the fact that normalization in the special case of subject HU's "loudness" data seemed rather to disguise than to further reveal the physical facts, emphasizes the necessity of further investigation in this area. One important problem is how to achieve running information about the exact distance between the object lens and the glottis. Kiritani (1971, 1972) has contributed to the solution of this problem by research on X-ray monitoring of the position of the fiberscope. The properties of the photo transducer is another fundamental problem which is further complicated by the lack of knowledge about the transportation of light below the glottis.

From the present results the following tentative conclusion might be drawn concerning the effect of "loudness" on the spatial control of the glottal gesture: repetitive productions of *p* with gradually increased "loudness" are accompanied by a systematic increase of the maximum glottal aperture. It cannot be generally stated whether there is a difference in the principle of control depending on the "loudness" being weak or not. It seems, however, to be a common feature for the subjects that the principle of spatial control is different when the "loudness" level is low and, for subject JR, perhaps also when the speech assumes a shouting character, but these assumptions must be supported by further investigation.

It is interesting that, within a certain range of "loudness" variation, a linear correlation can be observed between intensity (i.e. dB levels) and the size of the glottal aperture. This must lead to the hypothesis that the control of the degree of glottal abduction in connection with "loudness" variations takes a logarithmic scale as point of departure.

b. Speaking rate: For both subjects speaking rate and size of maximum glottogram amplitude also show a significant correlation. Figure 6 seems to display a close, approximately linear relationship between the degree of maximum abduction and segment duration all through the series.

Later in this paper the results of the spatial realization of the glottal gesture, for the "loudness" as well as for the rate series, will be reconsidered and incorporated as part of a larger framework together with the temporal relationships of the gesture.

B. THE EFFECT OF "LOUDNESS" AND SPEAKING RATE ON THE TEMPORAL RELATIONS OF THE GLOTTAL GESTURE

1. THE GLOTTAL ABDUCTION

The variation in the duration of the glottal abduction through the "loudness" and the rate series is shown graphically in figure 7.

a. "Loudness" It is clear from the figure that the range of variation in the duration of the glottal abduction is very small, about 10 ms for subject JR and less for subject HU. For subject HU the variation seems rather random, the correlation coefficient being zero, whereas the variation estimated from the correlation coefficient seems to display some systematism in relation to "loudness" for subject JR. However, the significant correlation seems to depend entirely upon the values of the peripheral steps in the series; the correlation coefficient for the steps 2-7, the most interesting range as far as size of glottal aperture is concerned, is close to zero. Consequently, we may conclude that increased "loudness" has virtually no effect on the duration of the glottal abduction; irrespectively of the intensity of the test sound the glottal abduction lasts about 100 ms for subject JR and 120 ms for subject HU.

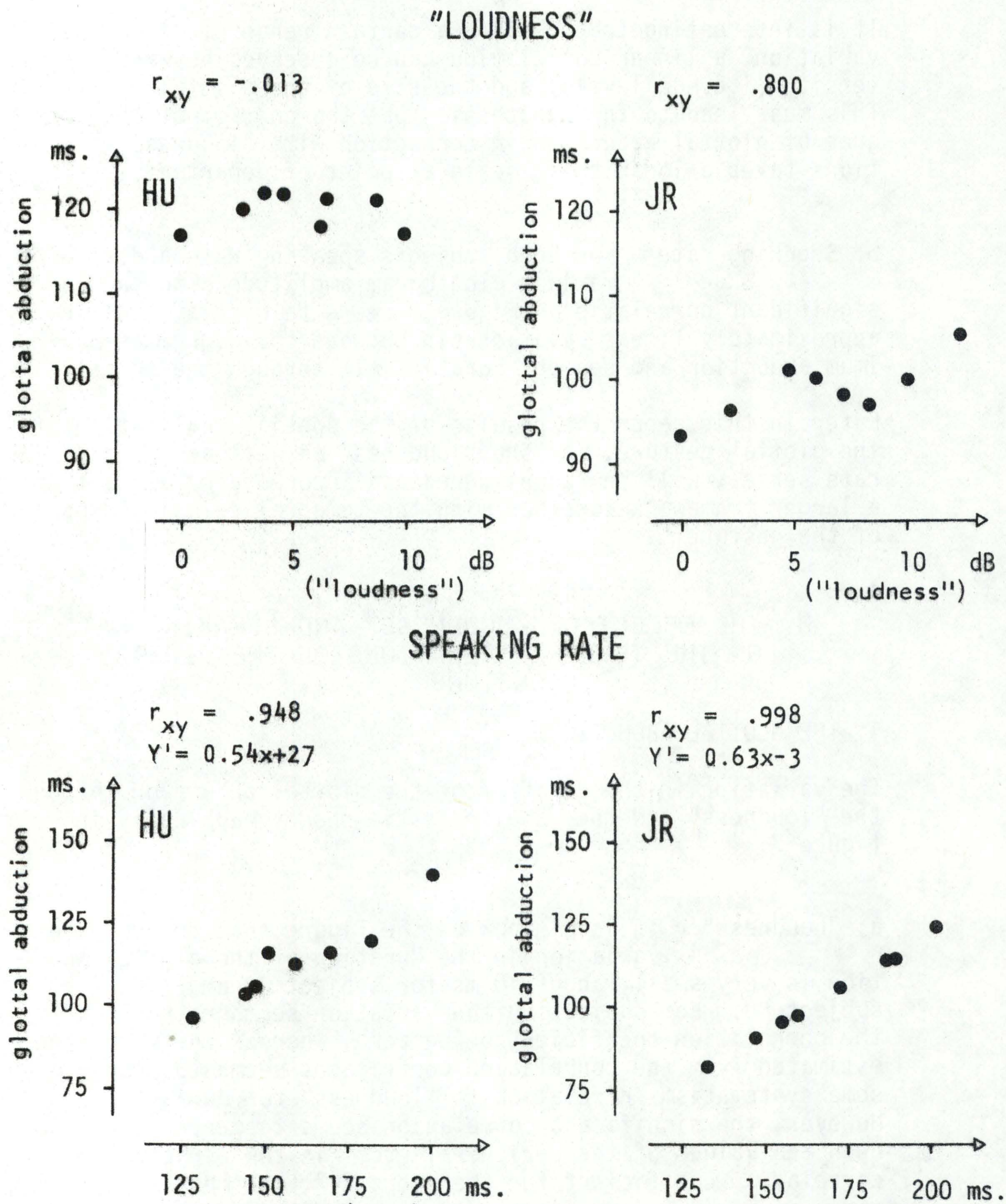


Figure 7

Duration of glottal abduction of *p* as a function of "loudness" (intensity of *p*) and speaking rate (duration of *p*).

b. Speaking rate: In connection with varying speaking rate the range of variation in the duration of the glottal abduction is much wider. Quite contrary to the relations at increased "loudness" - but not unexpectedly - the figure shows a clear, systematically shorter duration of glottal abduction concurrently with increased speaking tempo. The correlation between the segment duration and the duration of glottal abduction is linear, positive, and highly significant (1%).

Since the speaking rate of the test sound is represented by the segment duration, delimited on the acoustic curves, it can be concluded that a faster completion of the oral articulation of the test sound implies a proportionally (one-to-one) faster completion of its glottal abduction. The duration of the abduction is more reduced per step in the series with subject JR than with subject HU. The amount of reduction from slowest to fastest tempo in the series, predicted from the regression line, is on the average 45 ms for subject JR and about 38 ms for subject HU, which means that the duration of the glottal abduction is reduced by 55% through the series for subject JR and only about 40% for subject HU. For the sake of comparison, it is interesting to recall that the reduction of the segment duration from slowest to fastest speaking rate was 70 ms or 50% on the average for both subjects.

c. The relation between duration and degree of glottal abduction: In the previous section it was stated that there is a significant correlation between the degree of maximum glottal abduction and "loudness" and speaking rate, respectively, of the test sound. The duration of the glottal abduction was shown to be of constant value, irrespective of "loudness", but to correlate highly with speaking rate, the duration of the abduction being reduced proportionally with the segment duration. These relationships, or lack of same, might lead to interesting conclusions about the average rate of glottal abduction under each of the two extra-linguistically controlled speech conditions. As the intensity of the test sound (i.e. "loudness") has been shown to correlate highly with degree, but not with duration of glottal abduction, one might not expect to find a correlation between duration and degree of glottal abduction by increased "loudness". Contrariwise, as segment duration (i.e. speaking rate) correlates highly with both degree and duration of the glottal abduction, a high correlation between duration and degree of glottal abduction would be expected by increased speaking rate. These assumptions are confirmed by figure 8, which illustrates these correlations graphically. The two parameters, degree and duration of the glottal abduction, do not correlate in the "loudness" series but show for both subjects a highly significant (1%), positive, linear correlation in the rate series. As the degree and duration of the glottal abduction can be proved to decrease proportionally through the rate series, it might consequently be concluded that the average velocity of the abduction must be constant in the produc-

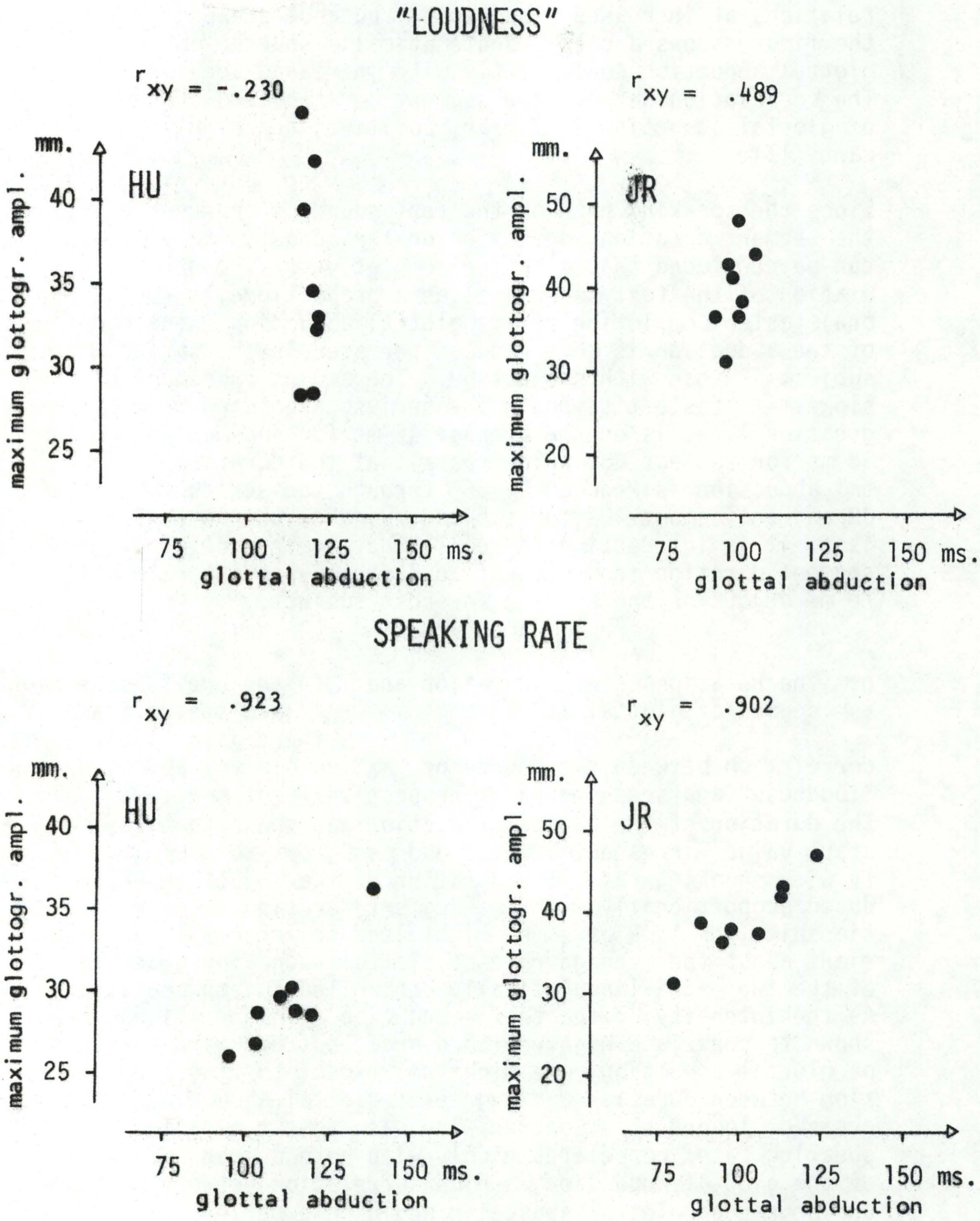


Figure 8

Maximum glottogram amplitude of *p* as a function of the duration of the glottal abduction. The relation is shown for the "loudness" series (upper half) and the rate series (lower half).

tion of the test sound repeated with increased speaking rate. Concerning the average of the glottal abduction in the production of the test sound repeated with increasing "loudness", it was previously stated that the maximum glottal abduction increases proportionally with "loudness" within a certain range of variation. As the time to reach the maximum abduction is constant, irrespective of "loudness", it must be concluded that the velocity of the glottal abduction is greater with increased "loudness".

2. THE GLOTTAL ADDUCTION

a. "Loudness" As shown in figure 9 the variation in the duration of the glottal adduction of the test sound is rather small and obviously random in relation to "loudness" for both subjects. The duration of the adduction averages 108 ms for subject HU, a bit shorter than that of the glottal abduction (120 ms), and 105 ms for subject JR, a little longer than that of the abduction (100 ms). On the whole, one might conclude about the relation between the duration of abduction and adduction that they each make up about 50% of the total duration of the gesture, irrespective of the degree of "loudness".

b. Speaking rate As opposed to "loudness", the effect of speaking rate on the duration of the glottal adduction seems rather different for the two subjects. Figure 9 indicates only a small effect of speaking rate on the duration of the glottal adduction for subject JR, i.e. only at very high and partly at very low rates does there seem to be a positive correlation, and this relationship is probably responsible for the significant (at the 5% level) correlation coefficient for the whole series, the correlation coefficient for the steps 1-7 is but .607, and for the steps 2-7 only .441. The marginal steps excepted, it can be concluded that the duration varies randomly about an average of 105 ms. For subject HU, however, there is a highly significant correlation between the two parameters and indications of a systematic reduction all through the series; the numerical reduction in the duration of the adduction is on the average 4-5 ms per rate step, which is a little less than that for the abduction. For this subject it is interesting to notice that the duration of the abduction and the adduction, respectively, expressed as percentage of the total duration of the gesture, makes up the same percentage, irrespective of speaking rate, namely 52-53% for the abduction and consequently 47-48% for the adduction.

During the adduction phase the onset of vocal cord vibration can be observed, from 20 up to about 75 ms ahead of the moment of completion of the adduction. The earliest start of oscillation is observed in the most forceful samples of the "loudness" series. This is not unexpected, as it is a well known fact that "loudness" implicates an increased transglottal flow,

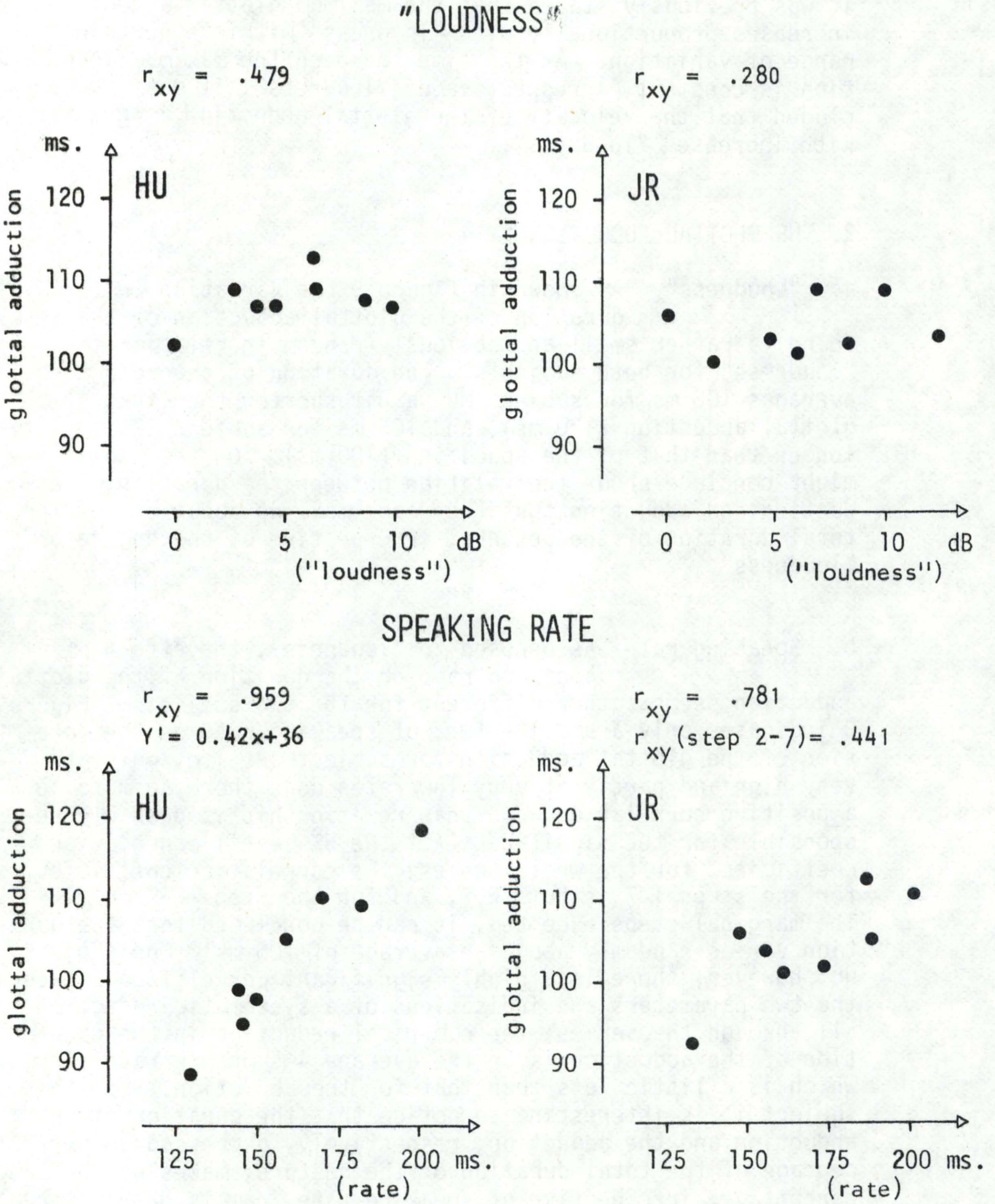


Figure 9

Duration of glottal adduction of *p* as a function of "loudness" (intensity of *p*) and speaking rate (duration of *p*) and the faster the flow after the release, the earlier the conditions for the Bernouilli effect to take place will be met. The highly significant correlation between extension of the vibration period through the adduction phase and "loudness" is illustrated in figure 10.

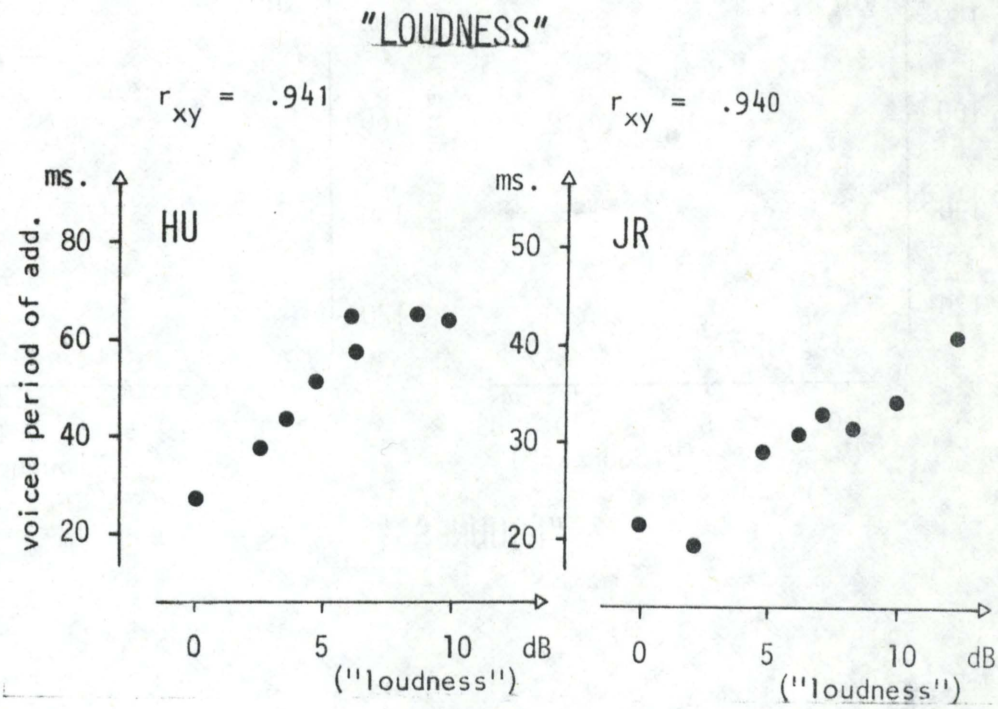


Figure 10

Duration of the voiced period during the glottal adduction of *p* as a function of "loudness" (intensity of *p*).

3. SEGMENT DURATION VS. THE TOTAL DURATION OF THE GLOTTAL GESTURE AND IMPLICATIONS OF THEIR DIFFERENCE FOR DELIMITATION CRITERIA

In the previous section it was stated that the duration of the glottal abduction and adduction, respectively, remained largely unchanged, irrespective of "loudness". Accordingly, it is not surprising that the total duration of the glottal gesture also remains largely unchanged through the "loudness" series, the duration averaging 225-230 ms for subject HU and 205 ms for subject JR. Figure 11 (upper half) shows clearly that segment duration, on the other hand, is systematically reduced when "loudness" is increased. This systematic reduction is almost exclusively due to the criteria used for delimitation of the aspiration of the stop and the onset of the following vowel. Defining the moment of oscillation onset as the start of the following vowel might actually lead to fatal generalizations

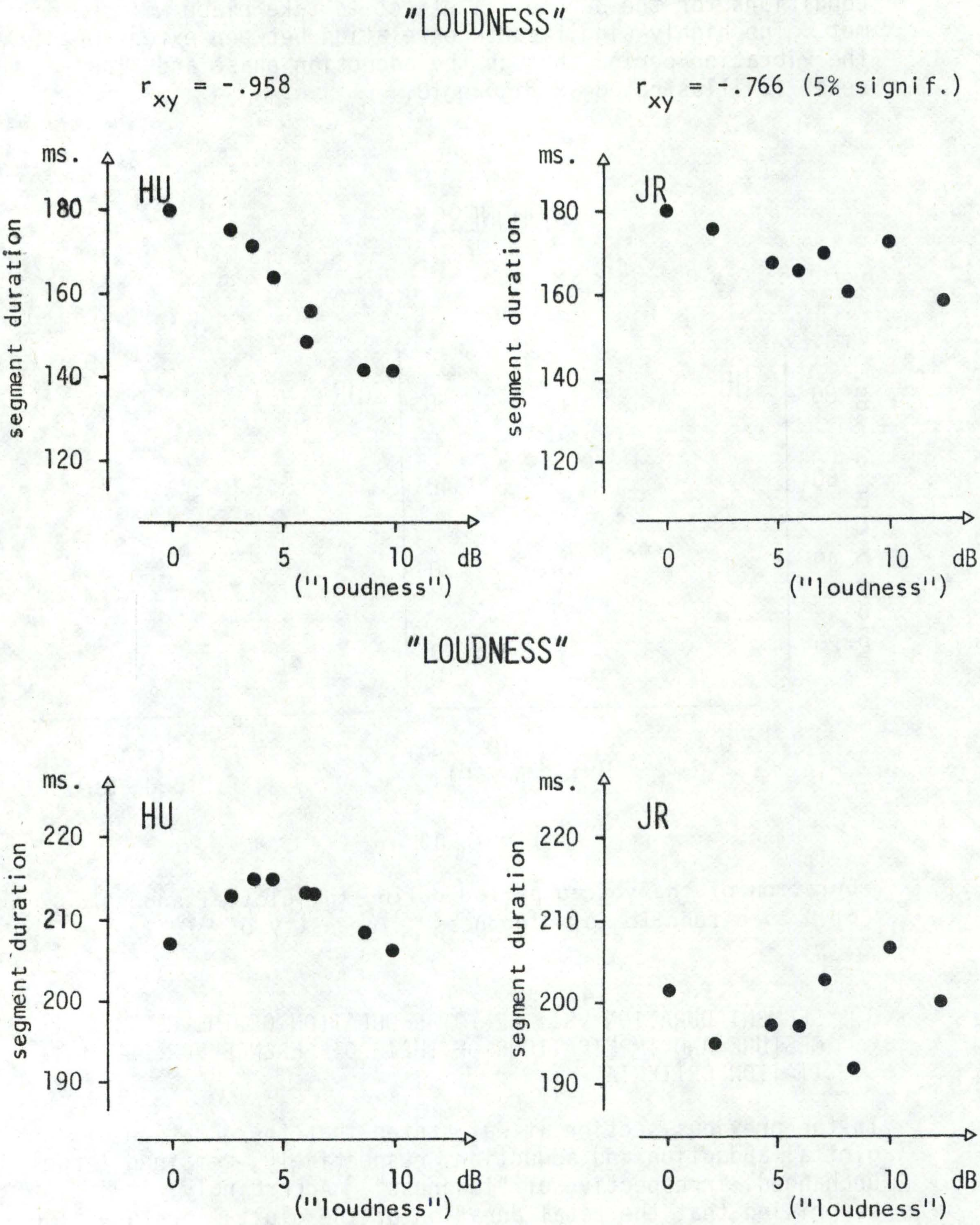


Figure 11

Duration of p as a function of "loudness" (intensity of p). In the upper half of the figure the relation is shown for duration of p being defined as the interval from oral implosion to the onset of vocal cord vibration. In the lower half the relation is shown for duration of p being defined as the interval from oral implosion to the moment of completion of the glottal adduction.

about the temporal organization of speech from the point of view of production, as the ever earlier onset of vocal cord vibration during the adductory phase at increased "loudness" may be considered to be a passive consequence of aerodynamic conditions (i.e. heightened subglottal pressure) rather than a programmed reduction in the duration of the aspiration and thereby in the duration of the stop. Seen from that point of view, the chosen criterion for delimitation implies ambiguities of up to ±40 ms for average consonant and vowel durations (in individual cases up to ±60 ms) if "loudness" is not taken into consideration.

An alternative delimitation criterion might have been preferred. If, for the moment, the start of the following vowel is defined as the moment of completion of the glottal adduction, then no systematic reduction in segment duration will occur (cf. figure 11, lower half). Using this criterion for delimitation makes the aspiration of the stop consist of a voiceless plus a voiced interval, where the voiceless interval takes up between 80% and 40% (depending on the degree of "loudness") of the total duration. The moment of completion of the glottal adduction, as determined by means of a physiological curve, seems to appear in the acoustic curves as the moment when the complete formant pattern of the following vowel has been stabilized (as regards the presence of upper formants and their intensity). For a more thorough treatment of the problems of delimitation, the reader is referred to Fischer-Jørgensen and Hutters (1981).

As one of the major purposes of the present paper is to shed light on the temporal organization of speech, it seemed preferable to redefine the concept of segment duration and, in the following, it will be defined as the period from the labial implosion to the moment of completion of the glottal gesture. It shall be emphasized that application of the redefined segment duration as the rate parameter would not appreciably have changed the results and conclusions reported so far concerning the effect of speaking rate.

C. SUPRAGLOTTAL AND GLOTTAL TIMING

This section describes the temporal relationship between oral closure and aspiration, glottal ab- and adduction, and the interarticulatory timing of the glottal and supraglottal events under the two speech conditions.

1. SUPRAGLOTTAL TIMING

It has been stated previously that increased "loudness" has no effect on segment duration, while increased speaking rate is characterized by a systematic, linear reduction in segment duration. Concerning general timing strategies of the closure and aspiration components under the two speech conditions, it might be preliminarily assumed that the shorter segment dura-

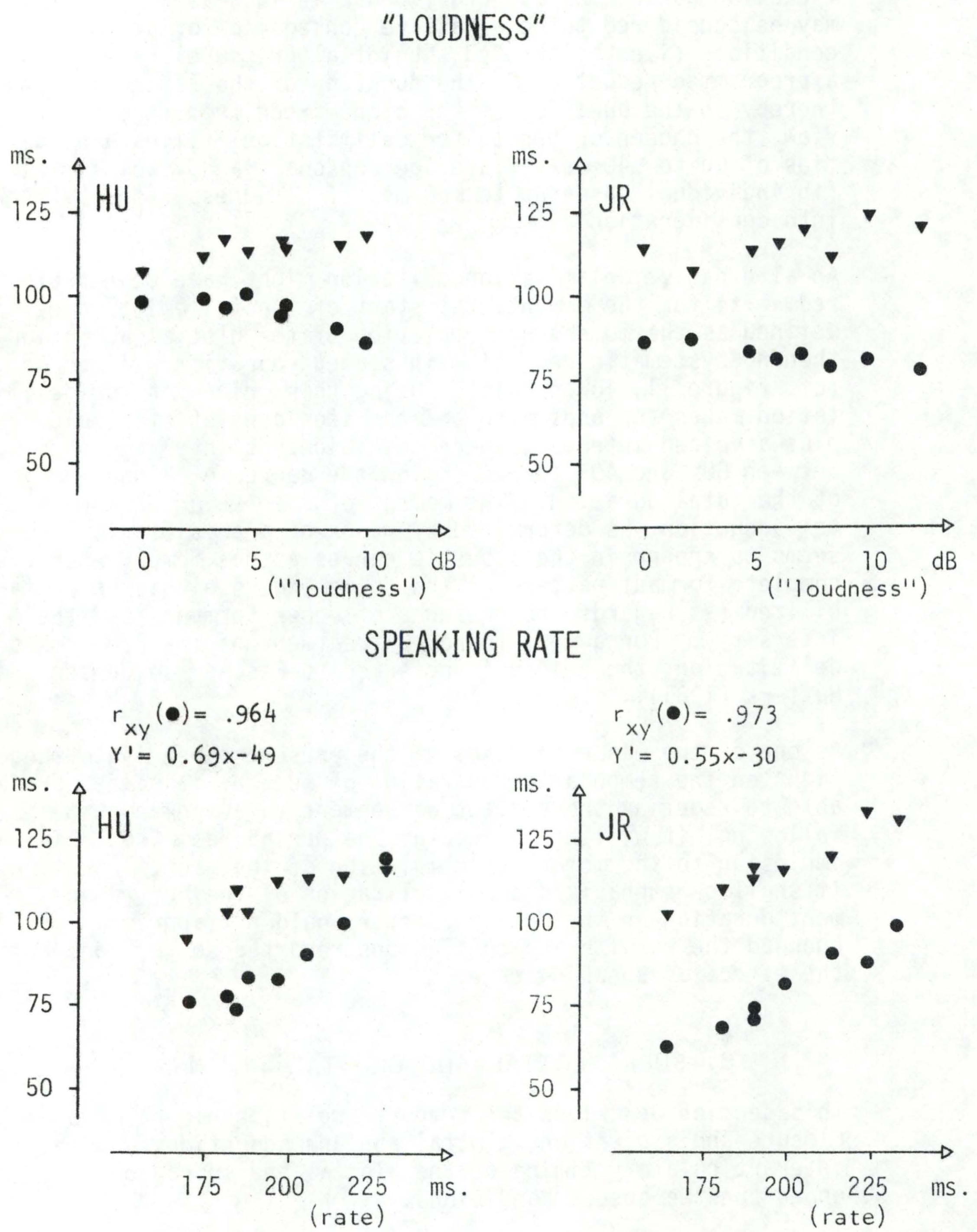


Figure 12

Closure duration (the dots) and duration of aspiration (the triangles) of *p* as a function of "loudness" (intensity of *p*) and speaking rate (duration of *p*) tion at increased rate is a consequence of reduction in the duration of the closure as well as the duration of the aspiration. As for the "loudness" condition, variation (if any) in the duration of closure and aspiration must be inversely correlated.

Figure 12 confirms that increased speaking rate involves a systematic reduction of both closure and aspiration. Though far less radical or systematic, the effect of increased "loudness" is a shorter closure as well but, as expected, the aspiration is prolonged. As for the rate condition, it seems to be a general strategy that the shorter duration is primarily a consequence of a reduction of closure duration, the tendency being more radical for subject HU, where closure reduction accounts for almost 70% of the total segment reduction; for subject JR the figure is only about 55%.

It appears from figure 12a that, for subject HU, shorter segment duration is primarily reflected in closure duration also under different "loudness" conditions. Accordingly, it seems as if the two subjects have different individual timing strate-

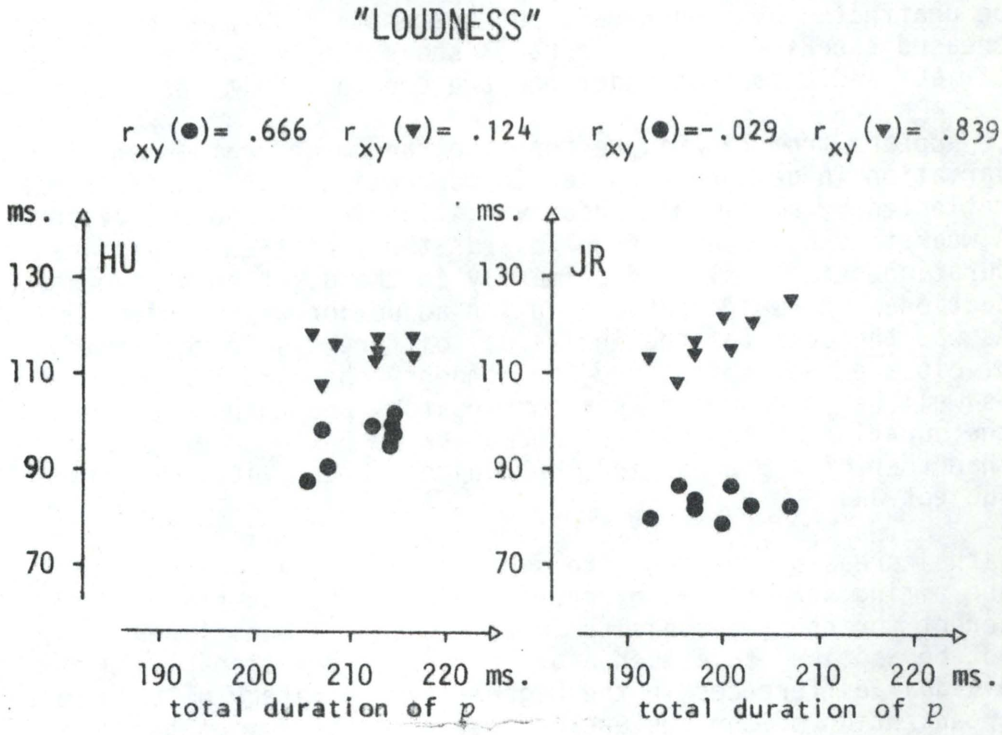


Figure 12a

Closure duration (the dots) and duration of aspiration (the triangles) of *p* as a function of the total duration of *p*. The relations are shown for the "loudness" series.

gies applied to the common constraints laid upon the temporal organization of closure and aspiration under the two different speech conditions. This might be explained as a mere compensatory phenomenon, since closure duration makes up a greater percentage of the segment duration for subject HU than for subject JR, i.e. subject JR has in general a relatively short closure, less susceptible to reduction than that of subject HU. The tendency to compensation is clearly illustrated by the fact that the reduction in closure duration from slow to fast rate amounted to 45 ms for subject HU and 38 ms for subject JR, while reduction in aspiration was 23 ms for subject HU and 30 ms for subject JR. However, it is interesting to observe that the reduction in duration and aspiration, expressed as a percentage of the duration at the slowest speaking rate, is almost identical for the two subjects, namely about 37.5% for the closure; for the aspiration it is 19.5% for subject HU and 22.5% for subject JR.

2. GLOTTAL TIMING

Like the segment duration (intended to express the supraglottal timing), the duration of the glottal gesture, in the abductory as well as in the adductory phase, was shown by and large to be unaffected by "loudness", but systematically reduced at increased speaking rate. Figure 13 shows the timing of the glottal ab- and adduction under the two speech conditions.

It appears from the figure that the random and rather small variation in gesture duration in connection with "loudness" is reflected by small and random variation in ab- and adduction. A weak tendency can be seen towards the variation in gesture duration being reflected primarily in the duration of the abduction with subject JR and in the adduction with subject HU. As was the case for the individual differences in the timing of closure and aspiration, the tendency observed here might as well be interpreted as a compensatory phenomenon, because the duration of the glottal abduction is on the average shorter than that of the adduction for subject HU and vice versa for subject JR.

With increased speaking rate it is clear from the figure that the timing strategy with regard to the glottal gesture is different for the two subjects in this case. The difference cannot be ascribed to compensatory processes but probably to individual differences in the higher-level strategy of temporal organization of the gesture. However, it is common to the two subjects that varying gesture duration is primarily reflected in the duration of the abduction. The primary difference in their timing strategy is that the shorter gesture duration seems to be a consequence of a systematic reduction of both ab- and adduction with subject HU, but to a much higher degree reflected exclusively in a systematically reduced abduction duration with subject JR. From the slowest to the fastest speaking rate the amount of reduction in abduction duration

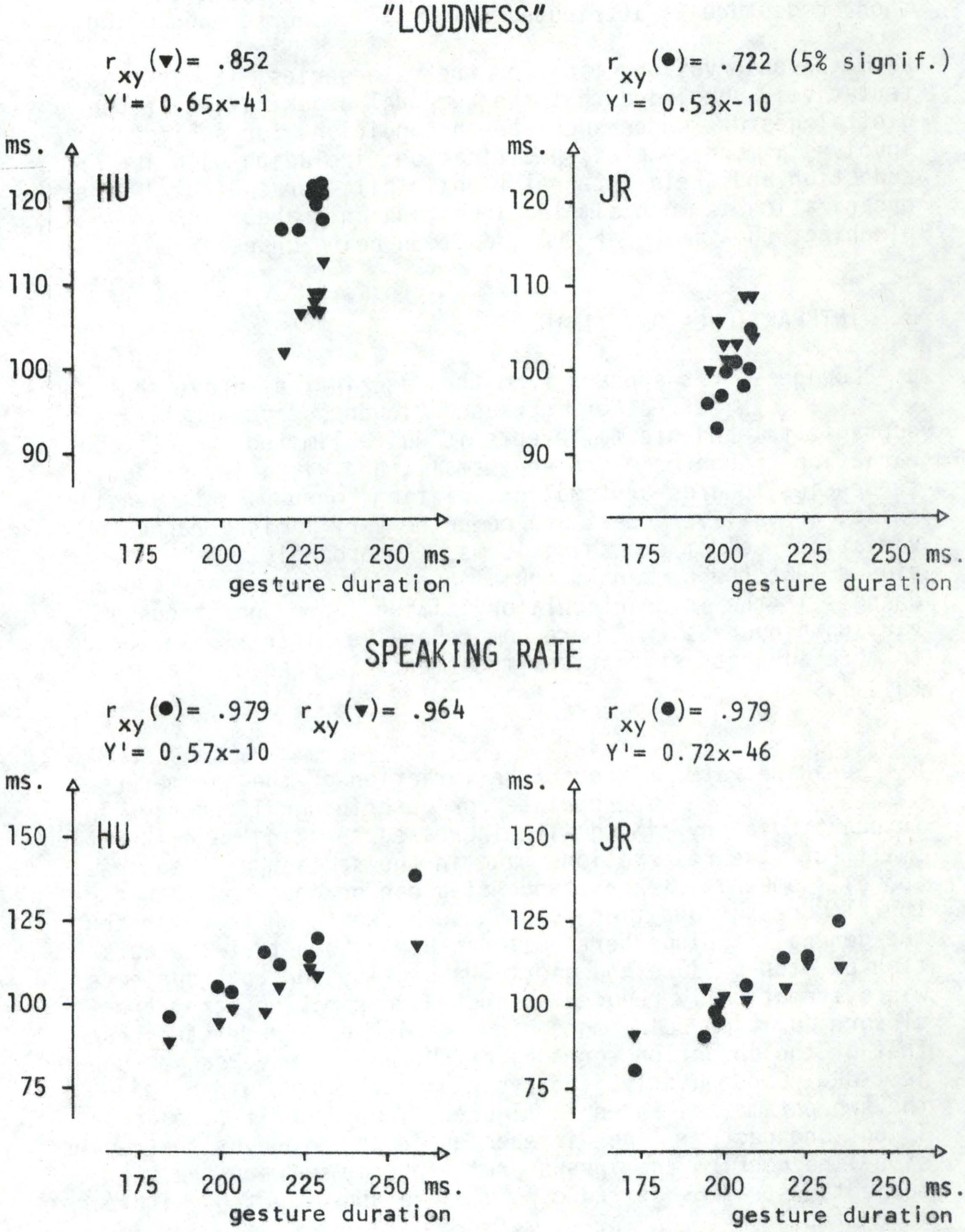


Figure 13

Duration of glottal abduction (the dots) and adduction (the triangles) of *p* as a function of the total duration of the glottal gesture. The relations are shown for the "loudness" series (upper half) and the rate series (lower half).

constitutes about 57% of the total reduction in gesture duration for subject HU and about 72% for subject JR. If the marginal steps (1,2 and 8) are omitted with subject JR, the shorter gesture is attributed solely to a shorter abduction.

From the observations made for the rate series, it might be tentatively concluded that the temporal organization of the glottal gesture under such speech conditions for subject HU involves a more complex specification, including both ab- and adduction and their interrelations, while for subject JR the organization is more simple, involving only abduction in the planning, the timing of the adduction being unspecified.

3. INTERARTICULATORY TIMING

a. "Loudness" It appears from the descriptions above that the effect of increased "loudness" on the timing of supraglottal and glottal events is quite limited. The temporal variations found here are very small, and virtually random; tendencies towards systematic variation, when such tendencies exist, are all very weak and comprise very limited ranges of variation, seldom exceeding 10 ms. Accordingly, it may be concluded that the present material does not suggest any clear changes in the interarticulatory timing under conditions of varying "loudness". A more comprehensive material, subjected to more advanced statistical treatment, is needed to clarify this issue.

b. Speaking rate Although the variation of the parameters appropriate for describing differences in interarticulatory timing with increased speaking rate is rather small (cf. the reservations made in the section on "loudness" above), somewhat clearer tendencies can be observed than for the "loudness" condition, and it seems worth-while to present the general findings here. It has been shown that the duration of stop closure and glottal abduction for both subjects was systematically reduced at increased speaking rate, the closure duration being more reduced with each rate step than that of the abduction for subject HU, and vice versa for subject JR. Consequently, the systematic relationship between the two parameters shown in figure 14, upper half (the dots), is not unexpected. The differences in the slope of the regression line confirm the inverse relationship between the two subjects' strategy of reducing closure and abduction with each rate step.

From the relationships established one might expect to find a change in the timing of the explosion and the moment of maximum abduction at increased speaking rate, the time span between the two parameters being increased with subject HU and decreased with subject JR. Further, a priori assumptions can be made if the relationship between closure duration and the period from oral implosion to the moment of maximum abduction

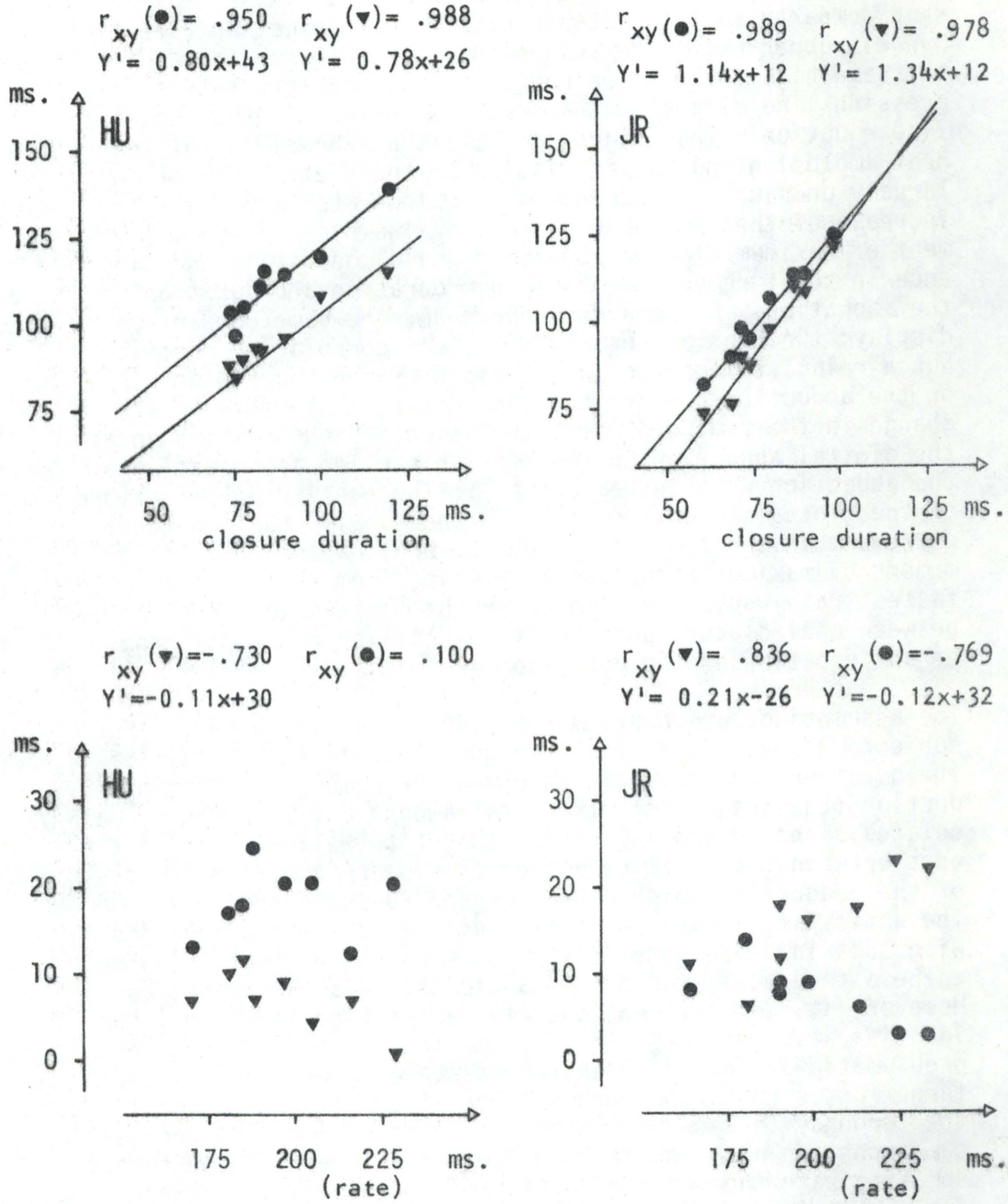


Figure 14

Upper half: Total duration of glottal abduction of *p* (the dots) and duration of glottal abduction of *p* with oral implosion as line-up point (the triangles) as a function of closure duration of *p* at increasing speaking rate. Lower half: The interval from the initiation of the glottal abduction of *p* to the oral implosion (the dots) and the interval from the oral explosion of *p* to the moment of maximum glottal abduction as a function of speaking rate (duration of *p*).

(i.e. duration of glottal abduction with oral implosion as line-up point) is taken into consideration. This relationship compared to that between closure and abduction (see figure 14, upper half, the triangles) shows for subject HU almost the same high correlation coefficients, nearly identical regression line slopes, but a smaller constant in the regression line equation. This might indicate that the timing of the oral implosion and the initiation of the glottal abduction is largely unchanged by increased speaking rate, and the total increase of the time span between oral explosion and the moment of maximum glottal abduction would constitute the difference in total reduction of closure duration and duration of the abduction. As regards subject JR, the same comparison displays two high and nearly identical correlation coefficients but a relative increase in the slope of the regression line on the abduction data points, indicating a tendency towards a change in the timing of oral implosion and the initiation of the glottal abduction in terms of an earlier anticipation of the abduction with increased rate. The earlier anticipation of the glottal abduction would further imply that the total decrement in the time span from the oral explosion to the moment of maximum abduction when going from the slowest to the fastest rate would cover not only the difference in reduction between oral closure and abduction but also the total amount of anticipation of the abduction.

The assumptions are confirmed by figure 14, lower half. For subject HU there is a clear tendency towards an increase of the duration from oral explosion to the moment of maximum abduction (the triangles), the total amount of which can be calculated as being about 7.5 ms. This fits well with the difference between the total reduction of closure duration and that of the abduction, which was 47 ms and 39 ms, respectively. The anticipation data (the dots) display a zero correlation with speaking rate, though it must be admitted that there seems to be a tendency towards less anticipation at increased rate. However, it has been mentioned above (section II,E) that the identification of the initiation of the gesture was not unproblematic at faster speaking rate because of glottal vowel target undershoot, and errors of measurement might account for the tendency to less anticipation. With subject JR, the assumptions are also confirmed. From the slowest to the fastest speaking rate the total reduction of the time span between oral explosion and the moment of maximum glottal abduction averages 14.5 ms; the difference between the total reduction of oral closure duration and abduction added to the total increase of anticipation constitutes 14.5 ms as well.

IV. CONCLUSION AND DISCUSSION

From the results described in the previous section the following can be concluded about the effect of "loudness" and speaking rate on the glottal behaviour of Danish *p*:

The spatial realization is affected systematically by "loudness" as well as by speaking rate. An increase of "loudness" is accompanied by an increase of glottal abduction, at least for a specific range of "loudness", whereas an increase in speaking rate causes a decrease of glottal abduction.

As for the temporal organization, varying "loudness" seems to affect neither the duration of the glottal gesture, nor the timing of the O-C gesture. The abduction and adduction have the same duration, irrespective of degree of "loudness", each corresponding to an average of 50% of the total duration of the gesture, the abduction being a little longer than the adduction for one of the subjects, and vice versa for the other. Increased speaking rate is accompanied by a systematic reduction of gesture duration. The strategy for the timing of the gesture under different rate conditions seems to be different for the two subjects. Subject HU achieves a reduced gesture duration by an equal reduction of abduction and adduction, whereas subject JR mainly reduces the gesture duration by shortening the duration of the abduction; only in the margins of the rate variation, i.e. from slow rate to the lower end of normal, conversational rate and from the higher end of normal, conversational rate to very fast speaking rate does there seem to be a reduction in glottal adduction duration as well.

On the basis of the present data, a model of the glottal behaviour under the two speech conditions can be set up for each subject. The models are schematized in figure 15. From the two sets of models it might be deduced that the control of the glottal behaviour under the "loudness" conditions is exclusively concerned with the spatial realization of the gesture, while for the rate conditions it seems as if the control mechanism involves spatial as well as temporal organization. However, it has been shown above that the degree of glottal aperture under the rate conditions can be expressed as a function of the abduction duration. On the basis of this relationship the following hypothesis shall be advanced concerning strategies in the planning of the glottal gesture under the two speech conditions:

The strategy applied in the control of the gesture by increased "loudness" aims exclusively at a reorganization of the spatial realization, resulting in a wider glottal aperture at increased "loudness". As the timing of the gesture remains unchanged at increased "loudness", the rate of the glottal abduction is increased in order to achieve a greater glottal aperture. As for the rate condition, the strategy implies primarily a reorganization of the temporal realization, resulting in a reduced gesture duration at faster speaking rates. This result can be achieved in different ways by different subjects. The spatial realization of the gesture is controlled by the duration of the abduction; the O-C gesture is under these circumstances to be considered merely as a ballistic movement of constant velocity, the size of maximum glottal aperture being but a passive consequence of the timing of the ab- and adduction.

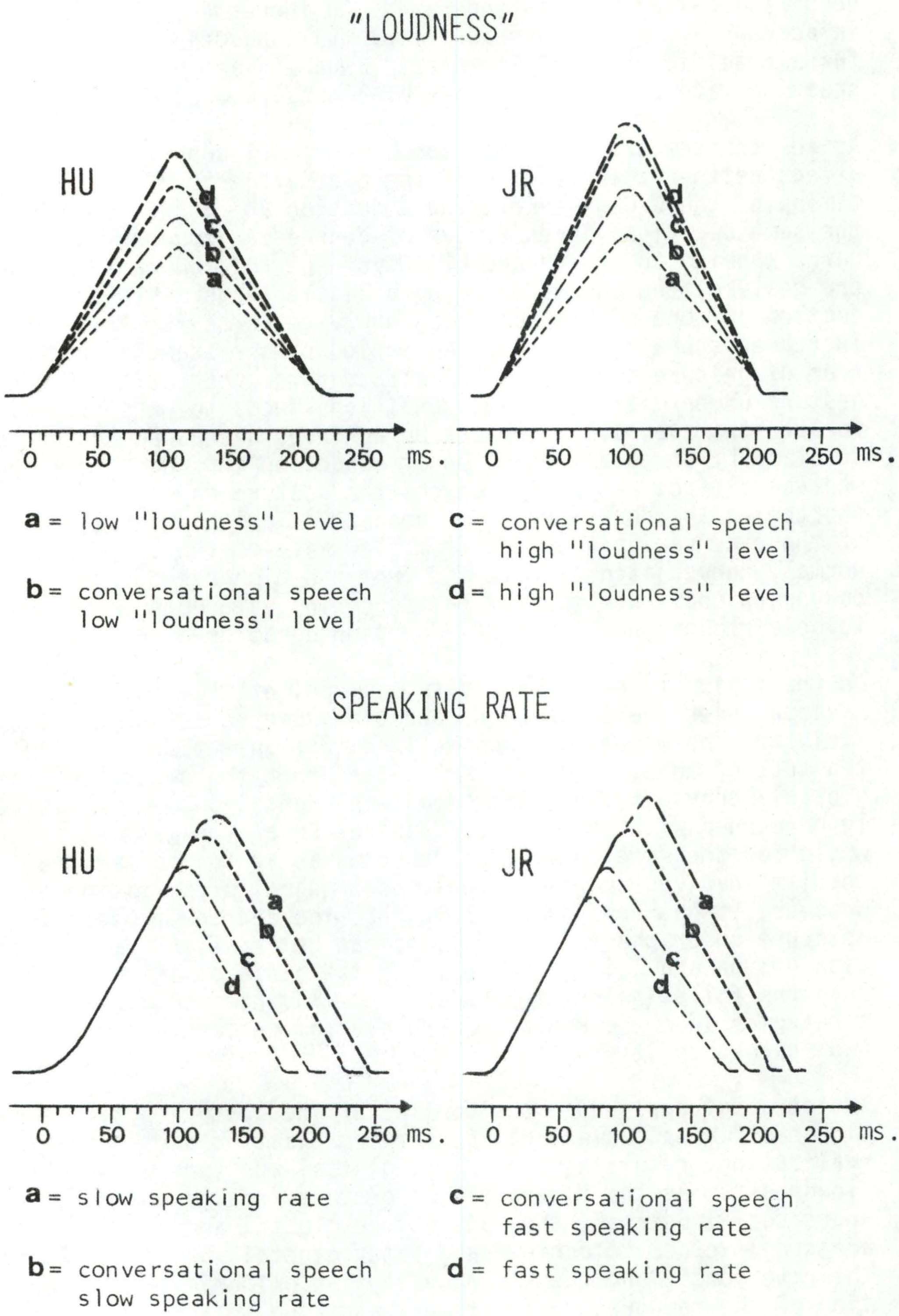


Figure 15

Schematized glottograms illustrating the O-C gesture of *p* at varying "loudness" and speaking rate (as for range of conversational speech cf. figure 4)

It is evident that the hypothesis advanced about the planning of the gesture must have interesting implications for the motor control of the glottal gesture. On the motoric level the increased rate of glottal abduction at increased "loudness" must presuppose an increase of the frequency of pulse firing to the posterior crico-arytenoid (henceforth PCA) muscle and/or a greater number of active motor units, i.e. it implies a change in the intensity but not in the timing of the motor commands to the abductor and probably the adductor muscles. At varying speaking rate the hypothesis presupposes an unchanged intensity of the motor commands to the PCA muscle and departs from the pattern of commands for the PCA-activity at slow rates. The relatively wide open glottis at the slowest rate must be a consequence of a rather strong PCA-activity, i.e. a high frequency pulse firing for the motor unit activation and/or a considerable number of active motor units; and an essential point in the hypothesis is that this motor program for PCA-behaviour is unchanged also at increased speaking rate, which means that the accumulated PCA-activity is the same and would have caused the same degree of glottal aperture as at the slowest rate; however, the size of the maximum aperture is decreased because the command to the interarytenoid muscle (henceforth INT) to an even higher degree will overlap the PCA-command and thereby the initiation of adduction will be triggered still earlier at faster tempo. Accordingly, the decrease of maximum aperture at increased tempo could be considered a mere consequence of the timing of the muscle command. As regards speaking rate control, the hypothesis assumes that the principles of Lindblom's (1963) model can be applied to the glottal behaviour of consonants as well.

It would have been useful to have EMG recordings of the laryngeal muscle activity of the present material in order to test at least that part of the hypothesis which is concerned with the motor control of the gesture. Unfortunately, only a very limited EMG material was available for this purpose. The material consisted of recordings from only one subject (HU) having spoken only a few repetitions of the phrase *"det er pile, de siger"* [d̥e: ˈb̥ʰi:lə d̥i ˈsi:ʌ] with increased "loudness" and speaking rate. However, the tendencies that could be extracted from the material give some support to the hypothesis advanced above. Figure 16 illustrates the interaction between PCA- and INT-activity of *p* under the two speech conditions. It shall be emphasized that the values on the vertical axes are all quite arbitrary. It appears from the figure that increased "loudness" obviously causes a reorganization of the motor program as regards the intensity of the commands but not their timing. At increased speaking rate the PCA-activity remains essentially unchanged. However, the decrease of glottal aperture through the rate series seems not to be exclusively a consequence of an increased overlapping of the PCA- and INT-commands, since the INT-activity, besides the change in timing, is also increased through the rate series. Accordingly, this aspect of the hypothesis is only partly supported by the results.

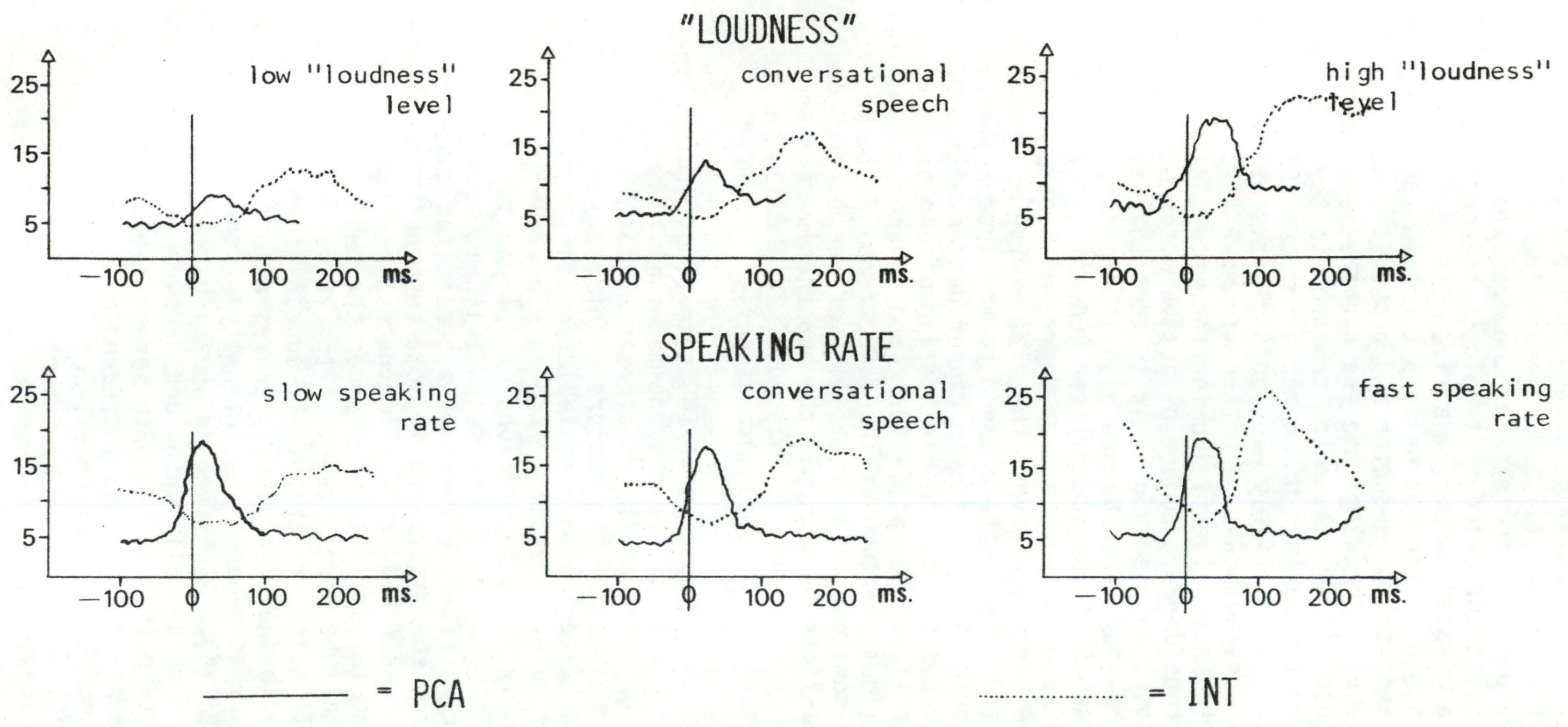


Figure 16

Tracings of PCA- and INT-activity of *p* in the phrase *det er pile, de siger* [ˌd̥e: ˈb̥ʰi:lə d̥i ˌsi:ɔ] at varying "loudness" levels and speaking rates (in this figure no distinction is made between different levels or rates of conversational speech). The implosion of *p* is used as line-up point. The scale used for the vertical axis is arbitrarily chosen.

The results of the present investigation do not give an unambiguous answer to the question raised in the introduction: whether an increase in speaking rate during stop consonant production implies a reorganization of glottal as well as supraglottal articulations in terms of faster rate of articulator movements and a change in the timing and intensity of the muscle commands. It appears, however, that the question of the existence of a common speaking rate control mechanism can be answered in the negative, if the production of a speech gesture is considered in terms of a coordinatory organization of two components: a movement towards and a movement away from a target, and if the comparison of supraglottal and glottal articulatory organization is made on this assumption. From the present investigation it appears that control of laryngeal behaviour implies an articulatory reorganization in terms of an increase or decrease in the rate of the glottal adduction only, while Gay et al. (1974) showed faster rate of both lip closing and lip opening gestures at increased speaking rate. At the muscular level it appears here as if only the command pattern of the muscle responsible for the movement away from the target is changed. In this connection, Gay et al. (1973) established a reorganization in timing as well as in intensity of the commands to the muscles responsible for both lip closure and lip opening gestures. If, on the other hand, the conclusion about the absence or presence of a reorganization of the driving force is to be based on changes in the amount of EMG activity of the total gesture (i.e. the movement both towards and away from the target), then the answer to the question is affirmative, as regards subject HU.

Since no EMG recordings of subject JR were included in the material, the question of PCA- and INT-interaction cannot be settled for this subject, but the different strategies employed by the two subjects in order to achieve a shorter gesture duration might indicate a difference between them in their organization of the laryngeal muscle activity. Whatever turns out to be subject JR's strategy of laryngeal control at increased speaking rate: one of pure timing changes, or one including also reorganization in terms of a decrease in muscle activity, neither of these alternatives corroborates the strategy of labial articulation proposed by Gay et al. (1974).

As mentioned in the introduction, it would not be surprising to reach a result like the one presented here, leaving the general impression that the supraglottal and glottal articulation of stops at increased speaking rate to a high degree work as two systems controlled by different mechanisms. Compared to the supraglottal gesture, the glottal gesture is assumed to be a simpler mechanism, less constrained as regards the necessity of reaching a fixed target, and with more degrees of freedom in its manner of contributing to the adequate aerodynamic characteristics of the stop.

Anyhow, most of the above considerations cannot be considered but as modestly supported hypotheses, and further investigations are needed, especially involving a more comprehensive study of the EMG-activity of the laryngeal muscles under different speaking rate conditions. Crucial points to be clarified seem to be: 1) Whether PCA-activity generally remains unchanged irrespective of the speaking rate (duration) of the stop. 2) The role of INT as a control mechanism by increased rate, i.e. can changes in the timing of the INT commands alone trigger the moment of maximum abduction, inferring that changes in the intensity of the commands have the separate function of regulating the rate of glottal adduction, or does an intention of earlier initiation of glottal adduction imply a change in timing as well as an increase of the INT-activity?

A problem connected to the questions posed above especially emphasizes the necessity of comprehensive EMG-studies. Generalizations about laryngeal control by varying speaking rate are often complicated, not only because different speakers might use different strategies, but perhaps even more by the fact that the different strategies are not necessarily different strategies of speaking rate control but can be assigned to general individual variations in laryngeal control. Sawashima et al. (1978) found individual variations in the sense that one subject was more dependent on the INT-muscle for laryngeal control, while the other relied more on the PCA-muscle. Bearing this in mind, it may not be safe to interpret the results presented concerning PCA-control of subject HU as being indicative of a special mode of laryngeal control at varying speaking rates, nor might we necessarily expect an EMG-recording of subject JR at increased speaking rate to display a PCA-behaviour equal to that of subject HU. Conclusions about strategies for speaking rate control with regard to laryngeal articulation presuppose a clarification of the general mode of laryngeal control for individual subjects.

V. NOTES

1. For example normalization of recording sections which display different linearity between glottogram amplitude and glottal aperture, the difference having arisen from alterations in the position of the fiberoptic cable caused by coughing, swallowing, or other reflex reactions of the subject.

ACKNOWLEDGEMENTS

The present paper is based on the author's thesis for the cand. phil. degree in phonetics. I wish to express my gratitude to Jørgen Rischel for his very constructive suggestions in connection with the present work. I am also indebted to Birgit Hutters and Peter Holtse for useful discussions, and to Svend-Erik Lystlund for technical assistance.

REFERENCES

Abramson, A.S., Lisker, L. and Cooper, F.C. 1965: "Laryngeal activity in stop consonants", *Status Rep. Speech Res., Haskins Labs. 4*, p. 6.1 - 6.13

Fischer-Jørgensen, E. and Hutters, B. 1981: "Aspirated stop consonants before low vowels, a problem of delimitation, its causes and consequences", *Ann. Rep. Inst. Phon. Univ. Cph., this issue*

Frøkjær-Jensen, B. 1967: "A photo-electric glottograph", *Ann. Rep. Inst. Phon. Univ. Cph. 2*, p. 5-19

Gay, T. and Hirose, H. 1973: "Effect of speaking rate on labial consonant production. A combined electromyographic/high-speed motion picture study", *Phonetica 27,1*, p. 44-56

Gay, T., Ushijima, H., Hirose, H. and Cooper, F.S. 1974: "Effect of speaking rate on labial consonant-vowel articulation", *J. Phonetics 2,1*, p. 47-63

Gay, T. 1978: "Physiological and acoustic correlates of perceived stress", *Language and Speech 21,4*, p. 347-353

Gay, T. 1978: "Effect of speaking rate on vowel formant movements", *J. Acoust. Soc. Am. 63*, p. 223-230

Harris, K.S., Gay, T., Sholes, G.N. and Lieberman, P. 1968: "Some stress effects on electromyographic measures of consonant articulation", *Status Rep. Speech Res., Haskins Labs. 13/14*, p. 137-152

Harris, K.S. 1971: "Vowel stress and articulatory reorganization", *Status Rep. Speech Res., Haskins Labs. 28*, p. 167-177

Harris, K.S. 1978: "Vowel duration change and its underlying physiological mechanisms", *Language and Speech 21,4*, p. 354-361

Kent, R. and Netsell, R. 1971: "Effects of stress contrasts on certain articulatory parameters", *Phonetica 24,1*, p. 23-44

Kiritani, S. 1971: "X-ray monitoring of the position of the fiberscope by means of computer controlled radiography", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics, Tokyo, 5*, p. 35-39

Kiritani, S. 1972: "X-ray monitoring of the position of the fiberscope in the pharynx", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics, Tokyo, 6*, p. 1-8

Kuehn, D.P. 1973: *A cinefluorographic investigation of articulatory velocities*, (Ph.D. thesis, University of Iowa)

Lindblom, B. 1963: "Spectrographic study of vowel reduction", *J. Acoust. Soc. Am. 34*, p. 1773-1781

Lisker, L. and Abramson, A.S. 1964: "A cross-language study of voicing in initial stops: acoustical measurements", *Word 20*, p. 384-422

Lubker, J.F. and Parris, P.J. 1970: "Simultaneous measurements of intraoral pressure, force of labial contact, and labial electromyographic activity during production of the stop consonant cognates /p/ and /b/", *J. Acoust. Soc. Am. 47,2*, p. 625-633

Öhman, S.E.G. 1967: "Word and sentence intonation: a quantitative model", *Speech Transm. Lab., Quart. Prog. and Status Rep., Royal Inst. Techn., Stockholm, 2-3*, p. 20-54

Sawashima, M., Hirose, H. and Yoshioka, H. 1978: "Abductor (PCA) and adductor (INT) muscles of the larynx in voiceless sound production", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics, Tokyo, 12*, p. 53-60

Sussman, H. and MacNeilage, P. 1978: "Motor unit correlates of stress: preliminary observations", *J. Acoust. Soc. Am. 64,1*, p. 338-340

Sussman, H. 1979: "Motor unit discharge patterns during speech: temporal reorganization due to coarticulatory and prosodic events", *Proc. 9 Int. Congr. Phon. Sc. II, Copenhagen*, p. 365-371