Consonant-dependent F0 variation

Katharina Pommerening and Merle Völkner

Institute of Phonetics and Digital Speech Processing, Christian-Albrechts University Kiel, Germany

The aims of this study are to investigate the effects of (1) the phonemic voicing distinction of German obstruents and (2) the allophonic variation between voiceless aspirated and voiceless unaspirated stops on the fundamental frequency (f0) and the open quotient (oq) at the onset of the following vowel. Furthermore, the influences of single voiced and voiceless consonants on these parameters were compared with the consonant cluster / ʃpr / and the affricate / t͡s /. The results indicate that the values for f0 and oq are higher at the onset of the following vowel after a voiceless obstruent than after a voiced obstruent. Aspirated stops cause even higher f0 and oq values. The results of the voiced and voiceless consonant clusters tend to be similar to the results of the single consonants.

1. Introduction

By varying the fundamental frequency speakers convey important linguistic and paralinguistic information to the listener. Beside the clearly intended suprasegmental f0-variation of the speaker there is also the segment conditioned variation of the fundamental frequency, e.g. the vowel and consonant specific f0-variations. Two hypotheses exist to explain the segment conditioned f0-variation: 1. The variation is the automatic consequence of the articulation. 2. The variation is intended by the speaker to supply additional acoustic cues to the listener which is called feature enhancement. In the current study these contrary assumptions will be tested on the basis of consonant inherent variation in f0 in the following vowel which depends on the voicing status of the preceding consonant.

Jessen (1999) found that f0- and F1-contours start at a higher level in vowels following an aspirated stop compared to an unaspirated stop. Moreover he describes the coherence between glottal opening and aspiration: A larger glottal opening causes a longer voice onset time, so that a longer aspiration duration arises which results in higher f0 and oq values after aspirated stops.

The relation between the intonation of a sentence and the intrinsic f0 values of obstruents is examined by Hanson (2001). Her results show that the f0 differences between voiced, voiceless and voiceless aspirated obstruents at the vowel onset are substantially higher in a stressed context than in an unstressed one.

The results of Löfqvist et al. (1988) are similar. In their study of cricothyroid muscle activity they point out that there is more tension in the vocal folds during a voiceless fricative than during its voiced counterpart. More tension is assumed to prevent vibration of the vocal folds and is achieved by a higher cricothyroid muscle activity. Increased vocal fold tension causes faster vibrations of the vocal folds at the onset of the following vowel which results in higher f0 values for the vowel following voiceless obstruents. They conclude that the f0-variation results automatically from articulation.

In contrast to this Kingston & Diehl (1994) take the view that the speaker varies f0 intentionally so that the listener recognizes phonemic distinctions more easily. They assume that speakers lower the f0 values by slackening of the vocal folds for voiced obstruents to enhance the contrast between voiceless and voiced stops (controlled speaking). In their view, this enhancement is linked with the phonological voicing status and not with the phonetic realization of the specific sound. Therefore they assume that the voiceless aspirated stop and the voiceless unaspirated stop following /s/ in clusters cause a similar f0 increase. Phonologically voiced stops which are often devoiced in languages such as English and German should induce a lower f0 than the phonologically voiceless stops. Therefore, according to the enhancement theory the phonologically voiced stop which is frequently realized as [b̥] and the phonologically voiceless stop following a sibilant in syllable onset which is realized as the unaspirated [p] should cause different f0 perturbations, even though from a phonetic point of view, they are produced similarly.

By investigating the f0 level between lenis and fortis plosives in different pitch patterns, rising, falling and monotone, Kohler (1982) also finds that f0 differences appear because of an active laryngeal control of the speaker.

The current study aims to shed new light on the question whether consonant intrinsic f0 variation is intended by the speaker or whether it is an automatic consequence of the voicing status of the consonant. First of all we provide new data concerning the studies by Löfqvist et al. (1988), Jessen (1999) and Hanson (2001). Secondly the influence of the consonant cluster / ʃpr / and the affricate / t͡s / is studied which has not been studied so far. We assume that the last consonant of the clusters have the greatest influence on f0. Thirdly, the glottal opening at the beginning of the following vowel is indirectly estimated by laryngographic data. The vocal folds are more tensed during a voiceless than during a voiced obstruent so that they vibrate faster in the following vowel after a voiceless obstruent. We revise the effects of this on the glottal opening phase.

The hypotheses of this study are the following:

(H1) For fricatives we should find that f0 starts higher after /f/ than after /v/.

(H2) If the f0 variation is more influenced by the aspiration than by the phonation status the f0 contour following /ʃp/ will be more similar to the ones following /b/ which is devoiced in German. If the phonological status would be crucial for the f0 contour following obstruents then the contour following /ʃp/ should resemble the one following the voiceless aspirated stop.

(H3) At the beginning of a following vowel after the affricate / t͡s / the f0 values are similar to the results of the single voiceless obstruents.

(H4) The f0 values after the consonant cluster / ʃpr / are similar to the results of the single voiced obstruents.

(H5) At the beginning of a following vowel the oq values are higher after an aspirated stop than after an unaspirated stop.

2. Method

2.1 Data Acquisition

Three male (*bma*, *fsa*, *jni*) and three female speakers (*awi*, *kro*, *mav*) of Standard German (all non-smokers between 22 and 24) were recorded by a dynamic microphone (Sennheiser MD 421) and a Laryngograph processor at 16 kHz sampling rate. The material consisted of nonsense words which were composed like this: [C (CC) V le]. The obstruents /b, p, f, v/, the affricate / t͡s / and the consonant clusters /ʃp, ʃpr/ were combined with the tense vowels /i, a, u/.

The first syllable of each test word was stressed and the accented test words were embedded in the carrier sentence "Ich habe ____ gesagt" (*I said ____.)*. Each sentence was repeated seven times in randomized order.

2.2 Measurements

All test words were acoustically labeled with the PRAAT software. The derivative of the Lx signal (DEGG) was also calculated by running a PRAAT script. f0 based on the DEGG signal was calculated by means of the software TKASSP, implemented by Michel Scheffers at the IPDS, Kiel. The open quotient was computed for all periods during the vowels using EMU and R (R Development Core Team 2004). The following formula is the basis of the oq calculations:

$$\text{oq} = \frac{\text{duration of the opening phase of the vocal folds}}{\text{duration of one period}}$$

The oq value shows the ratio between the duration of the opening phase of the vocal folds and the duration of one period. Therefore the oq values vary between 0 and 1. A higher oq value indicates a more breathy voice because of a longer glottal opening phase. A more modal voice results in a lower oq value. A threshold method was applied which used a 4/7 threshold for estimating the instant of glottal opening.

2.3 Statistical Analysis

Statistical analyses were carried out using the R software. ANOVAs were calculated for the main effects of vowel and consonant identity and their interactions. The parameters f0 and oq for the first and second period were the dependent variables. The vowels and the consonants were the independent variables. Pairwise t-tests with Bonferroni adjustments for multiple testing were calculated for significant effects.

The significant level is 5 % which has been lowered depending on the exact number of the comparisions since there were so many t-tests. You can find the relevant alpha value in the particular tableheadings.

3. Results

Table I shows the results of the calculated ANOVAs.

TABLE I: degrees of freedom, F-values and significance levels of the first and second measured point of time from f0 and oq for each speaker;
p> 0.05: n.s. (not significant), p<0.05: * (significant), p<0.01: ** (highly significant), p<0.001: *** (most significant)

awi	**df**	**f0 1**	**f02**	**oq 1**	**oq 2**	**kro**	**df**	**f0 1**	**f0 2**	**oq 1**	**oq 2**
a:, i:, u:	116					a:, i:, u:	113				
cons	6	28.5***	46.1***	9.5***	13.5***	cons	6	70.5***	67.7***	27.1***	17.0***
vow	2	85.3***	110.6***	20.1***	10.3***	vow	2	28.6***	35.0***	1.9	15.4***
cons*vow	12	2.3*	3.3***	1.4	1.8*	cons*vow	12	2.5**	2.9**	3.5***	5.1***
a:	36					a:	34				
cons	6	8.3***	10.3***	5.9***	13.4***	cons	6	18.5***	32.3***	15.9***	20.9***
i:	39					i:	38				
cons	6	12.5***	18.0***	1.8	3.2*	cons	6	19.3***	15.3***	9.0***	5.1
u:	41					u:	41				
cons	6	11.0***	24.4***	4.4**	3.1*	cons	6	32.4***	26.2***	10.5***	6.6
mav	**df**	**f0 1**	**f0 2**	**oq 1**	**oq 2**	**bma**	**df**	**f0 1**	**f0 2**	**oq 1**	**oq 2**
a:, i:, u:	123					a:, i:, u:	126				
cons	6	38.4***	37.8***	28.4***	23.3***	cons	6	140.3***	114.9***	39.6***	38.6***
vow	2	122.8***	132.2***	0.5	0.7	vow	2	242.3***	306.4***	19.3***	35.1***
cons*vow	12	2.7**	2.8**	6.9***	5.7***	cons*vow	12	8.2***	6.5***	32.5***	38.8***
a:	40					a:	42				
cons	6	9.3***	11.2***	34.1***	22.1***	cons	6	33.0***	31.1***	71.4***	87.2***
i:	42					i:	42				
cons	6	16.0***	14.8	8.2***	6.0***	cons	6	84.3***	52.1***	6.9***	8.9***
u:	41					u:	42				
cons	6	15.0***	14.8	5.3***	5.8***	cons	6	47.3***	41.5***	18.4***	15.2***
fsa	**df**	**f0 1**	**f0 2**	**oq 1**	**oq 2**	**jni**	**df**	**f0 1**	**f0 2**	**oq 1**	**oq 2**
a:, i:, u:	120					a:, i:, u:	124				
cons	6	25.2***	13.3***	24.2***	3.3**	cons	6	22.0***	40.1***	20.4***	18.2***
vow	2	34.1***	41.3***	14.9***	12.1***	vow	2	61.5***	193.3***	23.5***	16.9***
cons*vow	12	2.2*	0.1	6.0***	4.5***	cons*vow	12	4.0***	3.5***	13.5***	19.3***
a:	42					a:	40				
cons	6	12.3***	8.5***	36.0***	19.8***	cons	6	11.3***	25.6***	28.3***	43.8***
i:	42					i:	42				
cons	6	7.5***	6.8***	4.3**	0.9	cons	6	10.5***	20.3***	7.2***	2.7*
u:	36					u:	42				
cons	6	15.1***	3.7**	13.4***	2.9*	cons	6	8.2***	7.9***	3.3**	3.4**

(H1) At the beginning of a following vowel f0 starts higher after /f/ than after /v/.

The first hypothesis is confirmed for the mean values of all speakers. The f0 differences between /f/ and /v/ are significant for all subjects except for the following three cases: The vowels /a/ and /u/ of speaker *jni* and the vowel /i/ of speaker *fsa* do not show any significances (see table II).

TABLE II: Degrees of significant differences between the f0 values of /f/ and /v/; $p > 0.0125$: n.s., $p < 0.0125$: *, $p < 0.0025$: **, $p < 0.00025$: ***

	a	i	u
bma	.0000 ***	.0000 ***	.0000 ***
fsa	.0000 ***	n.s.	.0003 **
jni	n.s.	.0000 ***	n.s.
awi	.0032 *	.0004 **	.0005 **
kro	.0000 ***	.0000 ***	.0000 ***
mav	.0011 **	.0000 ***	.0000 ***

(H2) If the f0 variation is more influenced by the aspiration than by the phonation status the f0 contour following /ʃp/ will be more similar to the ones following /b/ which is devoiced in German. If the phonological status would be crucial for the f0 contour following obstruents then the contour following /ʃp/ should resemble the one following the voiceless aspirated stop.

All f0-contours after [p^h] show the highest values. In most cases the f0-contours of /ʃp/ are between the [p^h] and /b/ contours in the way it was expected (see figure 1).

There are some exceptions concerning the supposed order in which the f0 values after /ʃp/ are lower than after /b/ (speakers *fsa* and *kro*, vowel /i/), but these differences are not significant.

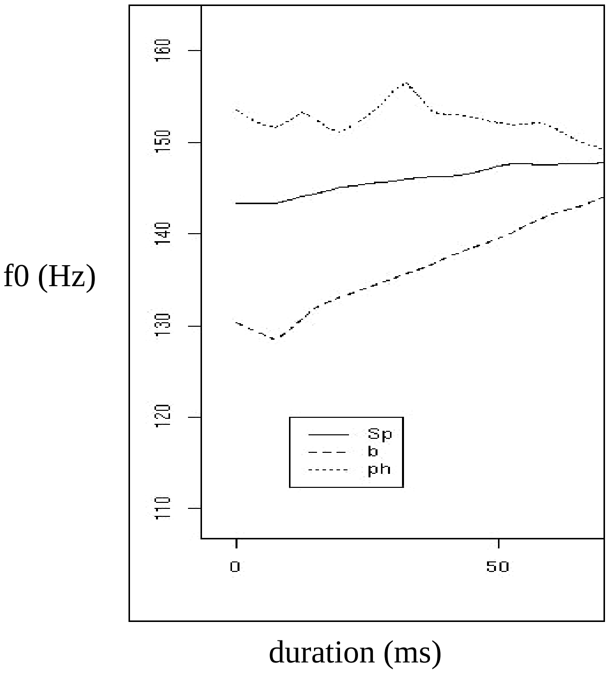

Figure 1: averaged f0-contours at the onset of the vowel /i/ following [p^h], /b/ and /ʃp/ of speaker *bma*

Table III shows the significant differences between [p^h] and /b/ for the first measured timepoint in the following vowel. The results are divided in two commensurate halves: There are nine cases which show significant differences between [p^h] and /b/ and nine cases which are not significant.

TABLE III: Significant levels for f0 values of [p^h] and /b/; $p > 0.0083$: n.s., $p < 0.0083$: *, $p < 0.0017$: **, $p < 0.00017$: ***

	a	i	u
bma	n.s.	.0000 ***	.0000 ***
fsa	n.s.	n.s.	.0000 ***
jni	n.s.	.0012 **	.0002 **
awi	.0072 *	.0010 **	n.s.
kro	n.s.	n.s.	.0000 ***
mav	n.s.	.0014 **	n.s.

There are no significant differences between the f0 values following [p^h] and /ʃp/ except in the following four cases: For the vowel /i/ the results of the speaker *bma* (**) show significance and for the vowel /u/ the results of the speakers *bma* (***), *fsa* (***) and *kro* (***) are significant (see table IV). The results are consistent because the values after /ʃp/ are always lower than after [p^h].

TABLE IV: Significant levels for f0 values of [p^h] and /ʃp/;
p > 0.01: n.s., p < 0.01: *, p < 0.002: **, p < 0.0002: ***

	a	i	u
bma	n.s.	.0004 **	.0000 ***
fsa	n.s.	n.s.	.0001 ***
jni	n.s.	n.s.	n.s.
awi	n.s.	n.s.	n.s.
kro	n.s.	n.s.	.0000 ***
mav	n.s.	n.s.	n.s.

The comparison between f0 following /b/ and /ʃp/ results in both significant and not significant differences. The majority of cases do not show any significances.The remaining five cases are significant: The f0 values for the vowel /a/ are significant for the speaker *jni* (***) and for the vowel /i/ the results of the speakers *bma* (***), *jni* (**) and *awi* (***) show significances. For the vowel /u/ the speaker *awi* (*) makes a significant difference between /b/ and /ʃp/ (see table V). The f0-contours after /ʃp/ show always higher values than after /b/ except of the contours of the speakers *fsa* and *kro* for the vowel /i/, as mentioned above.

TABLE V: Significant levels for f0 values of /b/ and /ʃp/;
p > 0.01: n.s., p < 0.01: *, p < 0.002: **, p < 0.0002: ***

	a	i	u
bma	n.s.	.0000 ***	n.s.
fsa	n.s.	n.s.	n.s.
jni	.0000 ***	.0009 **	n.s.
awi	n.s.	.0001 ***	.0084 *
kro	n.s.	n.s.	n.s.
mav	n.s.	n.s.	n.s.

Because the differences between /b/ and /ʃp/ were not significant in thirteen cases we were interested in the phonetic realization of the voiced stop. Therefore, the percentage of the devoiced [b̥]'s was calculated and shown in table VI. If there was no sustained periodic vibrancy in the Lx signal the /b/'s were characterized as devoiced. The male speakers (upper 3 contours) produced less frequently devoiced [b̥]´s than the female speakers.

TABLE VI: Instances of devoiced [b̥]´s
of the speakers in percent

bma	15
fsa	5
jni	10
awi	100
kro	95
mav	81

(H3) At the beginning of a following vowel after the affricate / t͡s / the f0 values are similar to the results of the single voiceless obstruents.

This hypothesis is confirmed. f0 following / t͡s / is nearly at the same level as f0 following /f/ and [p^h] (see figure 2). This is reflected by table VII and table VIII. There is only one significant difference between the f0 values of / t͡s / and /f/ for the vowel /i/ of the speaker *jni* (**). The differences between the f0 contours of / t͡s / and [p^h] are only in three cases significant: for the vowel /a/ of the male speakers *bma* (***), *fsa* (***) and *jni* (*). All the remaining cases are not significant.

TABLE VII: Significant levels for f0 values of / t͡s / and /f/; p > 0.017: n.s., p < 0.017: *, p < 0.003: **, p < 0.0003: ***

	a	i	u
bma	n.s.	n.s.	n.s.
fsa	n.s.	n.s.	n.s.
jni	n.s.	.0008 **	n.s.
awi	n.s.	n.s.	n.s.
kro	n.s.	n.s.	n.s.
mav	n.s.	n.s.	n.s.

TABLE VIII: Significant levels for f0 values of / t͡s / and [p^h]; p > 0.0125: n.s., p < 0.0125: *, p < 0.0025: **, p < 0.00025: ***

	a	i	u
bma	.0002 ***	n.s.	n.s.
fsa	.0008 **	n.s.	n.s.
jni	.0026 *	n.s.	n.s.
awi	n.s.	n.s.	n.s.
kro	n.s.	n.s.	n.s.
mav	n.s.	n.s.	n.s.

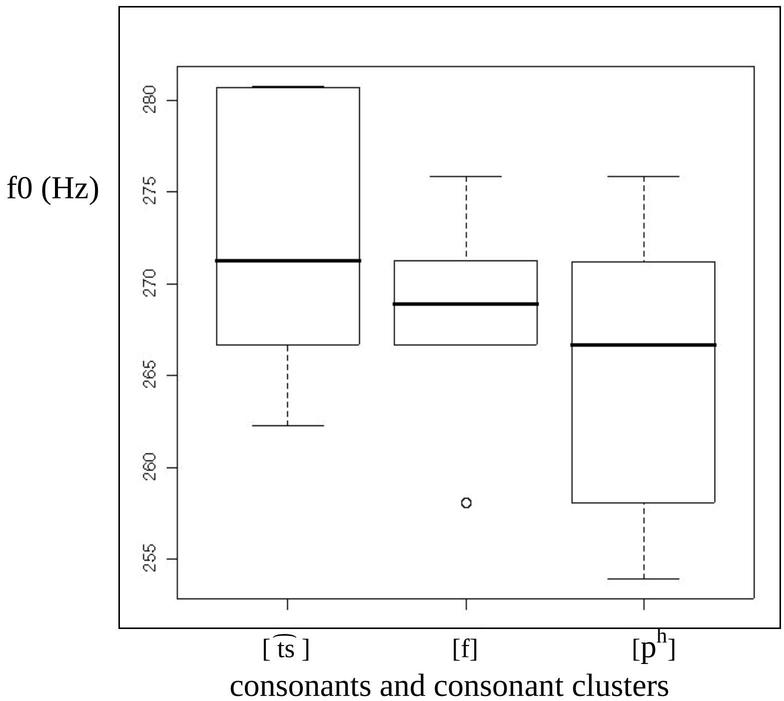

Figure 2: f0-boxplots at the starting point of the vowel /a/ for speaker *kro*; f0 in [Hz]

(H4) The f0 values after the consonant cluster /ʃpr/ are similar to the results of the single voiced obstruents.

The production of the /r/ in the sequence /ʃpr/ was inconsistent. In most cases the voicing status of the /r/ changed within the production from voiceless to voiced (see table IX, column "voiceless → voiced"). But there were also both fully voiced and voiceless productions of /r/.

In spite of these differences concerning the production of the /r/ the fourth hypothesis is verified (see figure 3).

TABLE IX: Instances of different productions of the phoneme /r/ in the sequence /ʃpr/ of the speakers in percent

	voiceless → voiced	voiced	voiceless
bma	71	-	29
fsa	79	5	16
jni	95	-	5
awi	95	5	-
kro	90	10	-
mav	61	29	10

The f0 values of /ʃpr/ and /v/ do not show any significances except of the following four cases: The results for the vowel /i/ are significant of the speakers *awi* (**) and *mav* (***) and for the vowel /u/ of the speakers *awi* (***) and *kro* (**) (see table X). There are also only four significant results for the comparison between the f0 values in the following vowel after /ʃpr/ and /b/. These results are presented in table XI. The significant cases occur for the vowel /i/ of the speakers *bma* (**), *awi* (*) and *mav* (**) and for the vowel /u/ of the speaker *fsa* (*).

TABLE X: Significant levels for f0 values of /ʃpr/ and /v/; $p > 0.017$: n.s., $p < 0.017$: *, $p < 0.003$: **, $p < 0.0003$: ***

	a	i	u
bma	n.s.	n.s.	n.s.
fsa	n.s.	n.s.	n.s.
jni	n.s.	n.s.	n.s.
awi	n.s.	.0006 **	.0000 ***
kro	n.s.	n.s.	.0005 **
mav	n.s.	.0000 ***	n.s.

TABLE XI: Significant levels for f0 values of /ʃpr/ and /b/; $p > 0.0125$: n.s., $p < 0.0125$: *, $p < 0.0025$: **, $p < 0.00025$: ***

	a	i	u
bma	n.s.	0.0005 **	n.s.
fsa	n.s.	n.s.	0.0026 *
jni	n.s.	n.s.	n.s.
awi	n.s.	0.0103 *	n.s.
kro	n.s.	n.s.	n.s.
mav	n.s.	0.0005 **	n.s.

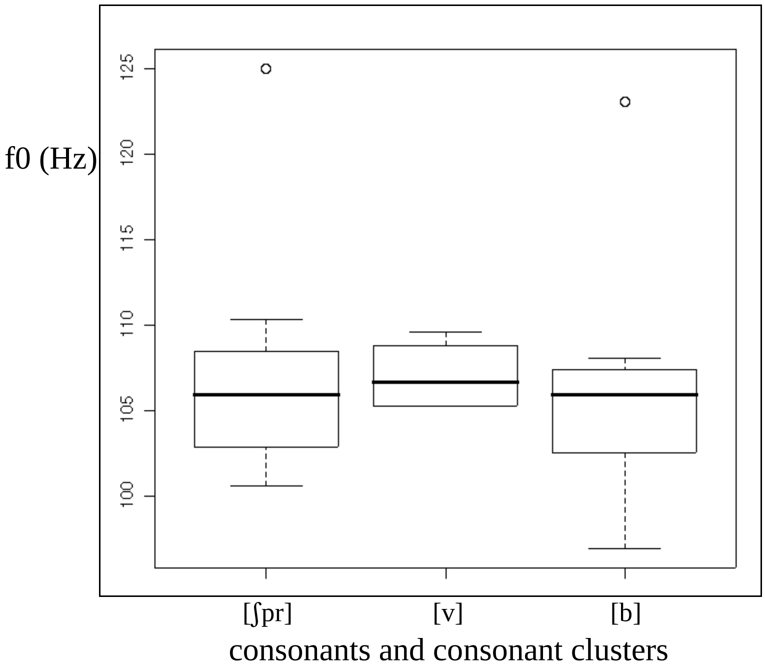

Figure 3: f0-boxplots at the starting point of the vowel /a/ of speaker jni; f0 in [Hz]

(H5) At the beginning of a following vowel the oq values are higher after an aspirated stop than after an unaspirated stop.

Since we assume that the oq values are only affected by the shape of the vocal fold vibrations but are independent of the shape of the vocal tract, the oq values for each subject were computed pooled for the three vowels.

For all speakers the oq-contours following [p^h] start at the highest level. For four subjects (*bma*, *fsa*, *kro*, *mav)* the contours following /ʃp/ have the lowest values. The contours of the vowels after /b/ start at values between them which is shown for speaker *mav* in figure 4. The order of the contours for /b/ and /ʃp/ of the other two speakers (*awi*, *jni*) was reversed (see as an example figure 5). However, none of the differences reached significance.

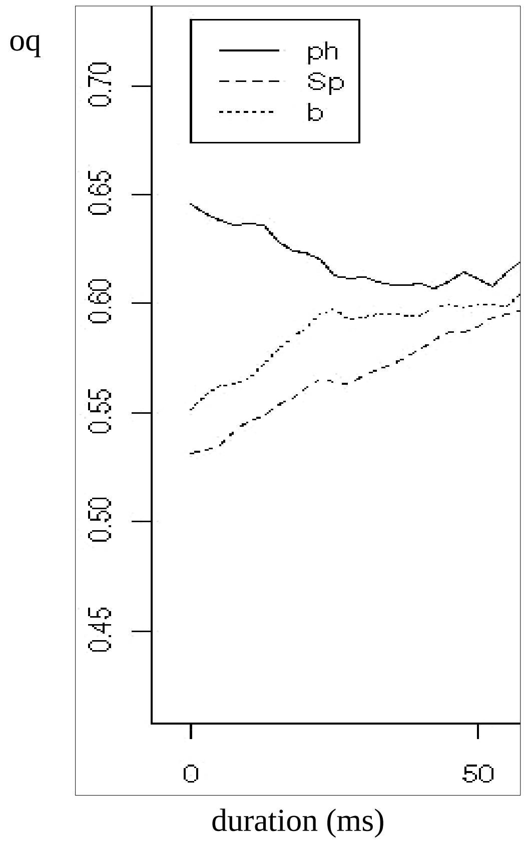

Figure 4: oq-contours for the three pooled vowels after [p^h], /b/ and /ʃp/ for speaker *mav*; axis of y: normalized values between 0 and 1

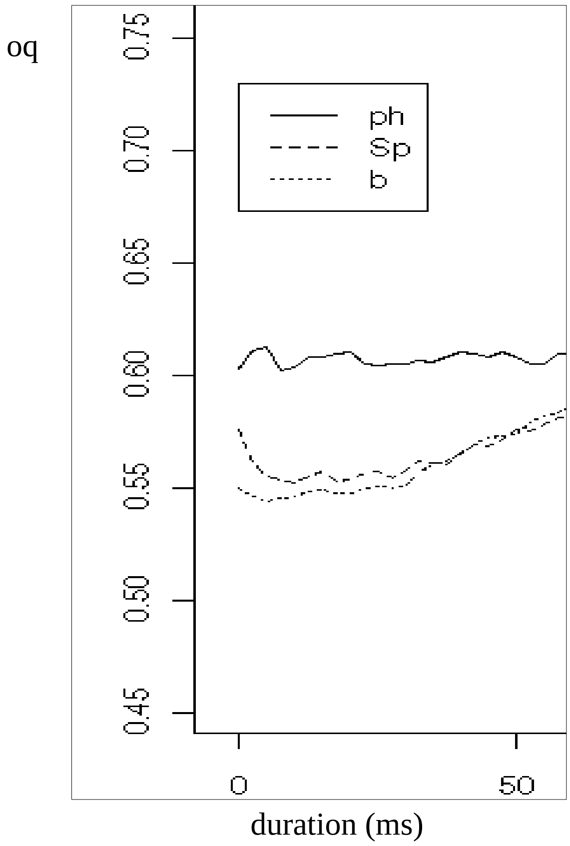

Figure 5: oq-contours for the three pooled vowels after [p^h], /ʃp/ and /b/ for speaker *awi*; axis of y: normalized values between 0 and 1

4. Discussion and Conclusions

On the one hand the aim of this study was to analyse the influence of single voiceless vs. single voiced obstruents and on the other hand to examine the influence of consonant clusters on f0 and oq of the following vowel. Moreover the question is discussed, whether different f0-values between voiceless and voiced obstruents result automatically from articulation or from feature enhancement intended by the speaker.

The first hypothesis is clearly confirmed for all speakers. The f0-contours after /f/ start always higher than after /v/. There are three cases in which the extent of the f0 difference between /f/ and /v/ is not as great as in the significant cases. This results in three not significant differences, but in spite of everything the f0 values of /f/ are higher than the values of /v/ in all cases.

The second hypothesis is confirmed. The fact that the f0 contours in the beginning of a vowel after the voiceless aspirated [p^h] start at a higher level than the f0 contours after the voiceless unaspirated /p/ in the /ʃp/-sequence can be explained by the stronger Bernoulli-effect. This effect occurs during an aspirated stop and causes higher f0 values because due to the increased airflow the vocal folds adduct more quickly.

Kingston and Diehl (1994) argue that a speaker tries to reduce the f0 contrast between two allophones of one phoneme. The results of this study do not verify this hypothesis because there are a lot of cases in which the f0 values of the /b/'s are similar to the f0 values in vowels following /ʃp/ (thirteen significant differences between /b/ and /ʃp/ and only five not significant differences), although they are not allophones of the same phoneme. These results can be explained by a closer inspection of the articulation of the voiced unaspirated stops. In almost all repetitions the female subjects produced the /b/ as devoiced [b̥]. The male speakers realized less frequently devoiced [b̥]'s, but there are devoiced [b̥]'s as well.

With regard to the results of the affricate / t͡s / the third hypothesis is confirmed. There is no different influence of the cluster on f0. The results show the same tendency as the results of the single voiceless obstruents because the obstruents in the cluster are voiceless, too.

Similar to the results of the / t͡s / sequence the f0 values after /ʃpr/ show the same tendency. It was expected that the consonant cluster /ʃpr/ would be produced voiced (hypothesis four). In nearly all cases the last phoneme of this sequence was voiced as the compared single consonant. The changing of the voicing status within the production of the /r/ results from the preceding voiceless consonant and the voiced vowel after the /r/, so that the /r/ is influenced by phenomena like assimilation and coarticulation. Inspite of the fact that the /r/ was sometimes produced from voiceless to voiced within the realization of the /r/ you can conclude that the f0 values after /ʃpr/ are similar to the f0 values after single voiced consonants. There are only a few cases of significant differences between /ʃpr/ and /v/ and between /ʃpr/ and /b/.

The fifth investigated hypothesis is verified for all six subjects. After an aspirated [p^h] the vocal folds are open for a longer time than after an unaspirated stop, as a result of more airflow through the glottis. If there is a preceding aspirated stop the beginning of a vowel is more breathy. Therefore the opening phase of the vocal folds is longer in the beginning of a vowel after aspirated stops than after unaspirated stops. That causes the highest oq values for the aspirated [p^h].

Generally the results of this study show that in a vowel following a voiceless obstruent the f0 values start higher than after its voiced counterpart. Some exceptions occur when /b/ was produced devoiced. In summary most of our results confirm the assumption that f0-differences develop automatically from articulation because it seems that the phonetic realisation is important for f0, but not the phonological voicing status.

References

Hanson H. (2001). Intrinsic effects of obstruent consonants on F0. *Journal of the acoustical society of America* 109: 2416

Jessen, M. (2001). Phonetic implementation of the distinctive auditory features [voice] and [tense] in stop consonants. In *Distinctive feature theory*, ed. T. Hall, 237-294. Berlin: Mouton DeGruyter.

Kingston, J., and R. L. Diehl. (1994). Phonetic knowledge. *Language* 70:419-454.

Kohler K.J. (1982). F0 in the production of lenis and fortis plosives. *Phonetica 39*:199-218.

Löfqvist A., Baer T., McGarr N., Seider Story R. (1988). The cricothyroid muscle in voicing control. *Journal of the acoustical society of America* 85:1314-1321

R Development Core Team (2004). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.