

---

# Effects of Frequency-Shifted Auditory Feedback on Fundamental Frequency of Long Stressed and Unstressed Syllables

Ulrich Natke  
Karl Theodor Kalveram  
Heinrich-Heine-University Düsseldorf  
Institute of Experimental Psychology  
Düsseldorf, Germany

---

Twenty-four normally speaking subjects had to utter the test word /tatatas/ with different stress patterns repeatedly. Auditory feedback was provided by headphones and was shifted downwards in frequency during randomly selected trials while the subjects were speaking the complete test word. If the first syllable was long stressed, fundamental frequency of the vowel significantly increased by 2 Hz (corresponding to 25.5 cents) under frequency-shifted auditory feedback of .5 octave downwards, whereas under a shift of one semitone downwards a trend of an increase could be observed. If the first syllable was unstressed, fundamental frequency remained unaffected. Regarding the second syllable, significant increases or a trend for an increase of fundamental frequency was found in both shifting conditions. Results indicate a negative feedback mechanism that controls the fundamental frequency via auditory feedback in speech production. However, within a syllable a response could be found only if the syllable duration was long enough. Compensation for frequency-shifted auditory feedback still is quite imperfect. It is concluded that control of fundamental frequency is rather important on a suprasegmental level.

**KEY WORDS:** fundamental frequency, pitch, frequency-shifted auditory feedback, auditory control, prosody

---

**D**eficits in the speech of postlingually deafened persons (Cowie, Douglas-Cowie, & Kerr, 1982) suggest that auditory feedback plays a significant role in speech production. The deterioration of speech takes place gradually; therefore, the role of auditory feedback for online corrections seems to be limited. However, altered auditory feedback can influence ongoing speech production. For example sidetone amplification decreases loudness (Siegel & Pick, 1974), and delayed auditory feedback prolongs vowel duration (Kalveram, 1984). The effects indicate that auditory feedback serves to control whether these parameters of an utterance match the intended values. A commonly used paradigm for investigating control loops is the acoustic perturbation technique, by which the original acoustic speech signal is artificially modified and then fed back via headphones. The direction of a response indicates whether it is a compensating mechanism. The latency as well as whether magnitude of the response is modulated by stimulus level are respective

indicators for a rather reflexive or voluntary response. The purpose of the present study was to examine the role of auditory feedback in controlling fundamental frequency during speech production. For this, the speech signal can be modified by an electronic device shifting all frequencies of the acoustic speech signal upwards or downwards. Thus, comparing the fundamental frequency produced by the speaker with and without frequency-shifted auditory feedback should then provide information about the control mechanism.

A number of studies using this paradigm indicate that a mechanism involving auditory feedback contributes to the control of fundamental frequency (Burnett, Freedland, & Larson, 1998; Elman, 1981; Kawahara, 1994; Larson, 1998). The experimental procedure used in these studies was as follows: Subjects had to produce a vowel for a duration of 5 s. After a randomly determined duration of 500 to 1500 ms following phonation onset, the frequency of the feedback signal provided via headphones was shifted electronically. Following this, the fundamental frequencies within a pre- and postshift interval were compared. The subjects were told to ignore the feedback changes and to maintain the same volume and pitch. In summary, the results showed a small compensation effect ("opposing responses"), but large interindividual differences occurred. Many subjects even showed a change of fundamental frequency in the direction of the frequency shift ("following responses"). Burnett, Senner, and Larson (1997) and Larson (1998) varied the frequency shift from 25 to 300 cents, where 100 cents equals one semitone. The fundamental frequency did not change proportionally to the frequency shift, but in the case of small shifts more opposing responses occurred, whereas in the case of large shifts more following responses resulted. Furthermore, the frequency shift obviously triggered two responses: The first response showed latencies between 100 and 150 ms, whereas the latencies of the second response were between 250 and 600 ms (Larson, 1998). A study of Burnett et al. (1998) revealed shorter latencies (192 ms) for the opposing responses than for the following response (327 ms). The authors concluded that small unexpected interferences will be compensated for by an automatic mechanism, whereas later responses represent a voluntary influence in order to adjust the fundamental frequency to be in keeping with an acoustic reference.

The effects of frequency-shifted auditory feedback have been studied in trained singers, too (Burnett et al., 1997; Parlitz & Bangert, 1998). When singing a musical scale or a straight tone, subjects were able to compensate completely for an unpredictable frequency shift of one semitone. Parlitz and Bangert (1998) reported that the first response in singers was a little faster than in nonsingers (mean latencies 113 ms vs. 135 ms),

and the second response with a latency of 300 ms was shown only by the singers. Furthermore, no relationship between latency and accuracy of compensation was observed.

The experimental procedure used by the cited authors revealed both immediate and short-term effects caused by frequency-shifted auditory feedback. Results showed the ability to respond to a frequency shift and to compensate for it especially while singing. However, these results are hardly applicable to natural speech, especially the production of syllables, because the experimental procedure referred only to continuous vocalization. Houde and Jordan (1998) applied formant-altered auditory feedback in whispering CVC syllables. The spectrum was altered in such a way that subjects heard different vowels than they actually produced. After an extensive training phase subjects compensated for the altered auditory feedback, and this compensation generalized to other vowels. Nevertheless, it remains unclear whether fundamental frequency is controlled online during speaking, especially within the production of syllables.

Regarding the control of vowel duration, a mechanism called "audiophonatory coupling," which works at the syllabic level, has been identified (Jäncke, 1991; Kalveram, 1984; Kalveram & Jäncke, 1989; Natke, 1999b). In the respective experiments, subjects had to utter a test word repeatedly while their auditory feedback was electronically either delayed or advanced in time during randomly selected trials. Under delayed auditory feedback vowel duration of long stressed syllables was prolonged up to 80% of the delay time. A delay time of 40 ms was used, which is not consistently perceived by the subjects (Jäncke, 1989). Vowel-related advanced auditory feedback was realized by playing back a vowel stored in a computer, which was triggered by the leading consonant of the respective syllable—for example, in a test word /tatatas/ (Kalveram, 1984). This led to an advanced feedback of the vowel about 30 ms (corresponding to the voice onset time) ahead of the actual onset of the vowel. In contrast with delayed auditory feedback, the advanced auditory feedback caused a shortening of the vowel duration of long stressed syllables. In both delayed and advanced auditory feedback the unstressed syllable remained almost unaffected. Additionally, in short stressed syllables the delay did not lead to a prolongation (Natke, 1999a). Results indicate that audiophonatory coupling makes sure that the planned vowel duration will be realized at least approximately, even in the case of varying vowel onsets, but it depends on the duration of syllables. In terms of control theory, the audiophonatory coupling operates as an open loop, feed forward mechanism preventing an erroneous vowel duration, at least in long stressed syllables (Kalveram, 1991).

In experiments involving the control of vowel duration the perturbation was applied intermittently. Subjects have to utter a test word repeatedly, and on single unpredictable productions the auditory feedback is modified. This technique prevents adaptation processes and makes sure that compensation mechanisms cannot be installed in advance. It is suitable to reveal speech control mechanisms that are likely to be involved in natural speech processes. This is especially true when the perturbation is not perceived by the subjects, because then the attention of the subject is not directed to the perturbation. The small perturbations then simulate small errors occurring during natural speaking. For example, these errors can be deviations of the intended vowel onset or of the intended fundamental frequency due to the complex neuro-physical processes involved in phonation.

The method applied in the present study is similar: perturbation of auditory feedback by intermittent application—now—of frequency shifting the auditory feedback while uttering a test word. This experimental procedure allows one to determine whether a mechanism for controlling fundamental frequency works in speech production and whether perturbations in the frequency domain can be compensated for within a syllable. As stated above, investigations regarding vowel duration revealed that delaying or advancing auditory feedback can be approximately compensated for within a syllable, given this syllable is long stressed (vowel duration about 320 ms). This is not the case in unstressed syllables (vowel duration about 140 ms) that seem to be too short for online control of vowel duration. Sapir, McClean, and Larson (1983) found responses in fundamental frequency 50 ms after auditory stimulation with clicks and interpreted them as responses of an auditory-laryngeal reflex. The latencies of the first responses under frequency-shifted auditory feedback were 100 to 150 ms (Larson, 1998). Therefore, an online compensation of a perturbation of fundamental frequency within a syllable seems to be possible—at least in long-stressed syllables. This was addressed in the experiment described below by using test words with different prosodic structures.

## Method

### Subjects

The study included 24 male German native speakers. None of the participating subjects showed a hearing deficit of more than 20 dB (audiometric test: Hortmann DA 323, Neckartenzlingen, Germany). Furthermore, the subjects did not show any neurological impairments nor did they take any medication that affects the nervous system, as reported in a standardized questionnaire. Because studies on the audiophony

coupling reported long-term adaptation effects even for measurement intervals of 4 weeks (Natke, 1999b), only those subjects were included who had not participated in an earlier experiment in which a test word had to be uttered repeatedly. The subjects did not reveal any speech or language disorder.

### Apparatus

The experiment took place in a sound-isolated chamber. The frequency shift was generated via a commercial device (DFS 404, Casa Futura Technologies, Boulder, CO) that works on a digital basis with a sampling frequency of 32 kHz and a sampling depth of 14 bits. This device was modified by installing a relay for switching between nonaltered auditory feedback (NAF) and frequency-shifted auditory feedback (FAF). A headset with closed ear pads (Blackhawk DSP 5DX, Flightcom, Portland, OR) was used, which itself attenuates the air conduction by 24 dB. For calibration the feedback gain was adjusted in a way that a sine-tone of 440 Hz and 75 dB (A) at the microphone led to a feedback volume of 70 dB (A) in the headphones. Subjects perceived the volume adjusted in this manner as a normal feedback volume. Binaural low frequency noise of 70 dB (A) produced by 900 Hz low pass filtered white noise was added during the whole experiment in order to mask bone conduction. In our experience low pass filtered noise masks the bone conduction much more effectively than white noise of the same dB level because bone conduction predominantly transmits low frequencies, whereas high frequencies mainly pass via air conduction. Because there is no simple way of physically measuring the effectiveness of the masking of bone conduction, we relied on the subjects' and our own observations. If subjects spoke with the masking noise but without auditory feedback of their speech signal, they reported that they did not hear their speech. If the speech signal was provided by headphones, subjects perceived this signal much louder than the masking noise, regardless of whether the speech signal was frequency shifted or not, and they reported that they did not hear their original speech signal when presented with the frequency-shifted signal.

An electroglottograph (EGG) (Laryngograph, Kay Elemetrics, Pine Brook, NJ) recorded the vibrations of the vocal folds. Control of the experiment and data acquisition was handled automatically by a commercial personal computer with a stereo soundcard. Recording was based on a sampling frequency of 11025 Hz and an amplitude resolution of 16 bits.

### Procedure

The subject's task was to speak the test word /tatatas/ with the speech rate and the stress pattern given by a

tone sequence presented twice via headphones before speaking. The target tones consisted of three sinusoidal tones of 440 Hz. An unstressed syllable was indicated by a duration of 200 ms, and a long stressed syllable by a duration of 400 ms. Two prosodic conditions were provided (/ta:tatas/ and /ta'ta:tas/). There was an interval of 3.5 s between tone sequences, during which the subject had to utter the test word. Presentation of the tones and recording were done automatically while the subject was sitting alone in the sound-isolated chamber. Subjects were asked to speak clearly and with a normal volume. No specific instructions regarding the fundamental frequency were given.

To get correct speech samples, in a training session subjects had to utter the test word simultaneously with the target tones presented regularly until the subjects produced the test word at least five times in succession with a correct stress pattern and syllable duration as judged by the experimenter. Generally, the subjects performed the training task correctly within 10 trials.

The experimental procedure included 30 trials stressing the first syllable and 30 trials stressing the second syllable of the test word. The other syllables had to be produced unstressed. From both stress conditions the fundamental frequency of stressed and unstressed syllables in first and second positions within the test word were available. The frequency of auditory feedback was shifted downwards in 20% of the trials. Frequency shifting was turned on after subjects had finished their last production of the test word and immediately before the next tone sequence. Frequency shifting did not modify the frequency of the tones. Subjects started to utter the test word while the device was already in frequency-shifting mode, so the whole test word was produced with frequency-shifted auditory feedback. After speaking the test word, frequency shifting was turned off immediately before the following tone sequence, and subjects heard their nonaltered voice in the next trial. The trials with frequency-shifted auditory feedback were selected randomly, but in a manner so that at least two trials with nonaltered auditory feedback preceded a frequency-shifted trial. The frequency was shifted by one semitone and .5 octave downwards in two groups ( $N = 12$  for each). The subjects were randomly assigned to these groups (shift of minus one semitone:  $M = 24.4$  years,  $SD = 2.76$ ; shift of  $-.5$  octave:  $M = 24.9$  years,  $SD = 2.86$ ). Prosody (i.e., stress pattern of the test word) varied within the groups.

These two shift magnitudes were used in order to produce (on the one hand) a small perturbation that might not be perceived by the subjects and (on the other hand) a greater perturbation that might lead to greater effects. After finishing the experimental procedure subjects were asked whether they had noticed an alteration

of the auditory feedback and, if so, whether they thought they had responded to this feedback alteration. The frequency shift of  $-.5$  octave was noticed by all subjects. Five of 12 subjects who experienced a shift of minus one semitone noticed an alteration of their speech feedback, but reported that they did not react to it.

## Data Analysis

Data analysis was performed using MATLAB (The MathWorks Inc., Natick, MA). The recorded speech signal served to check whether the test words had been spoken correctly. The EGG signal was digitally high pass filtered (Butterworth type) with a cut-off frequency of 50 Hz eliminating baseline variations. In order to determine vowel onset and offset the signal was rectified and smoothed by moving average, with 44 data points corresponding to nearly 4 ms. For a threshold for phonation activity a tenth of the signal's maximum value was selected. Because the test word contained only unvoiced consonants, the phonation equaled the vowel production. Fundamental frequency was determined using the high pass filtered EGG signal as follows. The glottal closing instant was characterized by a steep increase of the EGG signal (Childers & Krishnamurthy, 1984). This increase was assessed by determination of the maximum of the first derivative of the EGG signal. The period between two closing instants equaled the duration of one cycle, and its reciprocal equaled the momentary fundamental frequency. The fundamental frequency of a vowel, then, was calculated as the mean of these momentary values. At the beginning and at the end of the phonation there may have been an unstable phase of vocal fold vibration. Therefore the calculation of the mean fundamental frequency ( $F_0$ ) for a syllable was limited to the inner 80% of the complete phonation, whereas in unstressed syllables at least six cycles were included. For data analysis, fundamental frequency within each subject and each prosodic condition was averaged over the six trials, with frequency-shifted auditory feedback and the six nonaltered trials immediately preceding the shifted ones. There were no missing data.

For comparisons between frequency-shifted and nonaltered auditory feedback within the groups two-tailed  $t$  tests for dependent samples were calculated. Because four simultaneous tests were performed for each group the significance level  $\alpha = 5\%$  was corrected according to Bonferroni to  $\alpha' = \alpha/4 = 0.0125$ . Differences between groups were not tested statistically.

## Results

Table 1 and Figure 1 show the results for the frequency shift of minus one semitone. Figure 1 shows the

**Table 1.** Group results for the first and the second syllable of the test words /'ta:tatas/ and /ta'ta:tas/ for a frequency shift of one semitone downwards. Mean fundamental frequency  $F_0$  for nonaltered auditory feedback (NAF) and frequency-shifted auditory feedback (FAF) and difference ( $\Delta F_0 = F_{0,FAF} - F_{0,NAF}$ ) are shown in Hz. Standard deviations are given in parentheses. Statistics are based on two-tailed  $t$  tests for dependent samples ( $N = 12$ ).

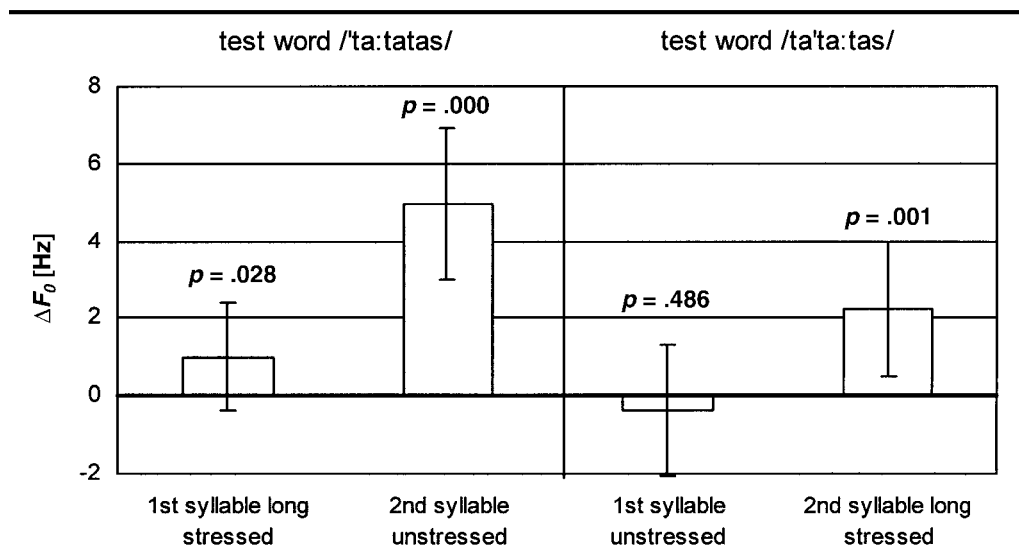
Test word and syllable	$F_{0,NAF}$	$F_{0,FAF}$ -100 cents	$\Delta F_0$	$t$ value	$p$ value	#subjects who showed increases	#subjects who showed decreases
/'ta:tatas/							
1 <sup>st</sup> syllable long stressed	128.4 (20.6)	129.4 (20.0)	1.0 (1.4)	-2.533	.028	10	2
2 <sup>nd</sup> syllable unstressed	128.8 (21.3)	133.7 (22.4)	4.9 (2.0)	-8.739	.000	12	0
/ta'ta:tas/							
1 <sup>st</sup> syllable unstressed	123.2 (19.4)	122.9 (18.5)	-0.3 (1.7)	0.721	.486	3	9
2 <sup>nd</sup> syllable long stressed	129.8 (22.1)	132.0 (22.0)	2.2 (1.7)	-4.479	.001	11	1

difference of the mean fundamental frequency ( $\Delta F_0$ ) between frequency-shifted auditory feedback and nonaltered auditory feedback. The left side of Figure 1 refers to the test word /'ta:tatas/ and shows that in the first long stressed syllable a trend is given for a small  $F_0$  increase of 1.0 Hz. This equals a mean response of 15.0 cents. Because the magnitude of the feedback shift is 100 cents, the response values in cents equal the percentage of the magnitude of the feedback shift. In the second unstressed syllable a significant increase of fundamental frequency by 4.9 Hz was found, which equals 65.2 cents. The right side of Figure 1 shows the results involving the test word /ta'ta:tas/. In the first unstressed syllable no effect was found. In the second long stressed syllable a significant increase of fundamental frequency by 2.2 Hz was found, which equals 30.9 cents. In Table 1

the mean fundamental frequency under each condition, the statistics, and the number of subjects who showed increases or decreases are presented.

Table 2 and Figure 2 show the results for the frequency shift of -.50 octave. Again the left side of Figure 2 shows the difference of mean fundamental frequency ( $\Delta F_0$ ) of the first two syllables of the test word /'ta:tatas/. In the first long stressed syllable the frequency shift resulted in a significant increase of 2.0 Hz. The response magnitude equals 25.5 cents, which is 4.25% of the magnitude of the feedback shift of 600 cents. The fundamental frequency of the second unstressed syllable increased significantly by 3.1 Hz. The response magnitude equals 44.9 cents, which is 7.48% of the feedback shift of -.5 octave. The right side of Figure 2 shows the results for the test word /ta'ta:tas/. No effect was found for the first

**Figure 1.** Differences of mean fundamental frequency ( $\Delta F_0$ ) between frequency-shifted auditory feedback with a frequency shift of one semitone downward (-100 cents) and nonaltered auditory feedback for the first and the second syllable of the test words /'ta:tatas/ and /ta'ta:tas/ ( $p$  values are based on two-tailed  $t$  tests for dependent samples).



**Table 2.** Group results for the first and the second syllable of the test words /'ta:tatas/ and /ta'ta:tas/ for a frequency shift of  $-0.5$  octave. Mean fundamental frequency  $F_0$  for nonaltered auditory feedback (NAF) and frequency-shifted auditory feedback (FAF) and difference ( $\Delta F_0 = F_{0,FAF} - F_{0,NAF}$ ) are shown in Hz. Standard deviations are given in parentheses. Statistics are based on two-tailed  $t$  tests for dependent samples ( $N = 12$ ).

Test word and syllable	$F_{0,NAF}$	$F_{0,FAF}$ -600 cents	$\Delta F_0$	$t$ value	$p$ value	#subjects who showed increases	#subjects who showed decreases
/'ta:tatas/							
1 <sup>st</sup> syllable long stressed	125.8 (17.5)	127.8 (18.6)	2.0 (1.6)	-4.195	.001	11	1
2 <sup>nd</sup> syllable unstressed	119.2 (18.3)	122.3 (18.2)	3.1 (2.5)	-4.258	.001	11	1
/ta'ta:tas/							
1 <sup>st</sup> syllable unstressed	122.7 (17.4)	122.9 (16.8)	0.2 (1.6)	-0.475	.644	5	7
2 <sup>nd</sup> syllable long stressed	122.4 (16.3)	125.2 (15.6)	2.8 (4.7)	-2.014	.069	8	4

unstressed syllable. In the second long stressed syllable there is a trend for an increase of the fundamental frequency of 2.8 Hz. The response magnitude equals 39.5 cents, which is 6.58% of the feedback shift. Table 2 presents the mean fundamental frequency under each condition, the statistics, and the number of subjects who showed increases or decreases.

As can be seen in the tables, interindividual variability in  $F_0$  is high. Nonetheless, in a repeated-measures design the effect may well be very strong if many of the subjects show a consistent reaction, as was the case here.

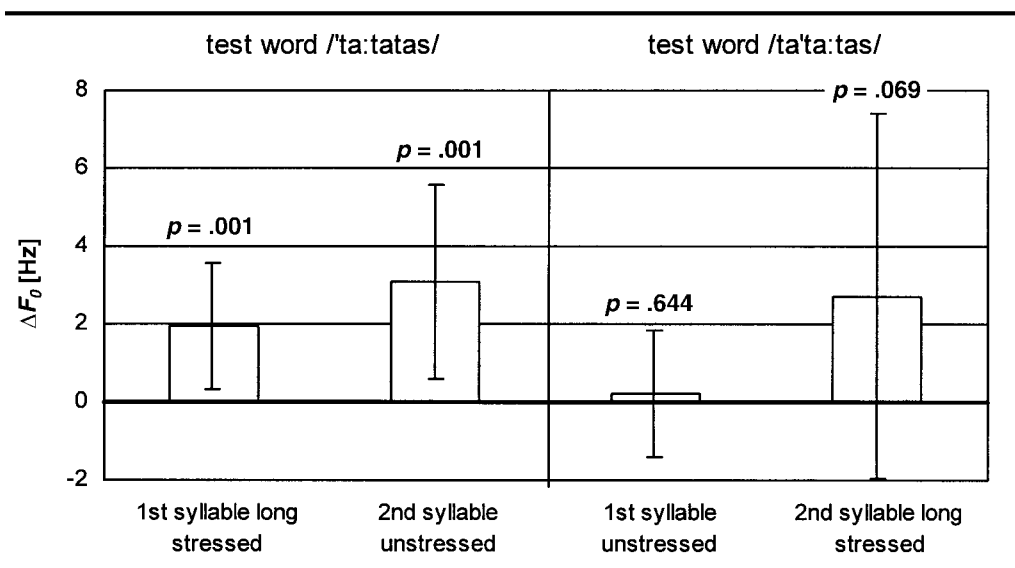
Additionally, vowel duration of the syllables was determined. Long stressed syllables had a mean vowel duration of 325 ms; unstressed syllables a vowel duration of 125 ms. Vowel duration of long stressed and unstressed syllables were not affected by the frequency-shifting

conditions, as are reported elsewhere (and for people who stutter as well) (Natke & Kalveram, 2001).

## Discussion

Frequency-shifted auditory feedback was found to change the fundamental frequency of the vowels of a test word. Even in the case of a small frequency shift, which was not noticed by the majority of subjects, a trend in the response was observed towards compensating for the frequency shift, though complete compensation was missed widely. The results suggest a compensatory negative feedback mechanism stabilizing fundamental frequency via the auditory feedback loop. Compared to earlier studies with continuous vocalization under frequency-shifted auditory feedback, the present investigation shows that

**Figure 2.** Differences of mean fundamental frequency ( $\Delta F_0$ ) between frequency-shifted auditory feedback with a frequency shift of  $-0.5$  octave ( $-600$  cents) and nonaltered auditory feedback for the first and the second syllable of the test words /'ta:tatas/ and /ta'ta:tas/ ( $p$  values are based on two-tailed  $t$  tests for dependent samples).



the effect is also found at the syllabic level in speech production. This suggests that the mechanism also applies to natural speech.

In this study frequency shifting was turned on before subjects started to utter the test word, so the complete word was produced under frequency-shifted auditory feedback. In the first syllable of the test word a response could be observed only when the syllable was long stressed. When the first syllable was unstressed and therefore short, no effect occurred. However, in the second syllable significant increases and a trend for an increase of fundamental frequency was found independently of whether it was long stressed or unstressed. The shortest latencies observed in studies with continuous vocalization are 100 to 150 ms (Larson, 1998). It can be assumed that long stressed syllables with a mean vowel duration of 325 ms are sufficiently long that fundamental frequency can be affected within these syllables, whereas latencies of the responses are too long to affect the fundamental frequency within unstressed syllables with a vowel duration of only 125 ms. These short syllables may be completed before the response begins. In the second syllable an effect may then be observed independently of its stress, because a response already occurred in the preceding long stressed syllable or the latency of the response equaled approximately the duration of the preceding unstressed syllable.

What do these findings mean for a mechanism controlling fundamental frequency in natural speech? If frequency shifting of auditory feedback is taken as a simulation of an error in speech production like a deviation of the intended pitch, one can conclude that the control mechanism is able to react to such an error within the ongoing syllable but only if it is long enough. Thus, errors in short syllables cannot be corrected within the syllable, but may lead to changes in following syllables. Therefore, the mechanism controlling fundamental frequency may lead to suprasegmental effects.

The magnitude of the responses observed in this study ranges from 15 to 65 cents. Compared to the frequency shift this is not a complete compensation, but it corresponds to the response magnitudes found in previous studies with continuous vocalization. Burnett et al. (1998) found a mean response magnitude of 29.2 cents ( $SD = 16.6$ ) when varying frequency shifts between 25 and 300 cents. Because in the present study effects of frequency shifts on mean fundamental frequency of vowels of syllables were examined, in the first syllables of the test word the fundamental frequency was averaged over the prereponse interval. The actual response magnitude in first syllables was likely greater than reported. In future studies it would be interesting to determine whether a response latency could be detected within the first syllable when looking at the fundamental frequency

over time and whether the peak response magnitude is greater than the magnitude reported here. In this study the data analysis was done automatically. Afterwards we analyzed the fundamental frequency over time in single syllables, which were long stressed and in first position of the word. Indeed we found some increases of the fundamental frequency approximately in the middle of the syllables. Systematic analysis of this type remains to be performed.

In this study, mainly opposing responses were found. Therefore the voluntary mechanism with greater latencies between 250 and 600 ms, as suggested by Larson (1998), which may adjust fundamental frequency to be in keeping with an acoustic reference (leading to following responses under frequency-shifted auditory feedback), seems not to operate in speech production on a segmental level. Results of this study suggest a negative feedback mechanism stabilizing fundamental frequency in speaking. Because compensation is incomplete, obviously the gain of this mechanism is low. In control of vowel duration (Kalveram & Jäncke, 1989), a similar effect regarding linguistic stress had been found: Changes in vowel duration under delayed or advanced auditory feedback were found only in long stressed syllables. It is suggested that vowel duration is controlled by a feed-forward process, which uses the temporal feedback of vowel onset to determine the time for terminating the vowel (Kalveram, 1991). However, stabilizing fundamental frequency resembles negative feedback control, even if compensation is missed. Considering both effects, it can be concluded that in the production of long stressed syllables peripheral feedback plays a greater role than in the production of short or unstressed syllables.

The question arises why the gain of the mechanism stabilizing fundamental frequency is low. The following answers are offered. First, a reference tone for regulating the fundamental frequency is missing when a person is not speaking in a habitual pitch. When a reference tone of 440 Hz, for example, is provided, trained singers can match it with an accuracy of more than 1 Hz (Sundberg, 1987). Furthermore, it was shown that an unpredictable frequency shift can be completely compensated for in singing (Burnett et al., 1997; Parlitz & Bangert, 1998). This indicates that the compensatory mechanism for fundamental frequency can be very effective if a reference tone is internally represented as in singing. In the case of speaking syllables, an intended value for the fundamental frequency of the syllable may also be given. However, in speaking, relative changes in fundamental frequency seem to be more important than keeping absolute values of pitch. Second, for precise comprehension of speech, fundamental frequency appears to be much less important than vowel duration, formants, and formant transitions. Thus, it seems less

important for the speaker to monitor and regulate fundamental frequency within the syllables. Consequently, although a compensatory mechanism for fundamental frequency at the syllabic level may exist, it seems that for languages such as English and German it is more important to control fundamental frequency at the suprasegmental level (e.g., for intonation). It would be interesting to employ the frequency-shifting method with speakers of languages such as Chinese, where pitch contour within a syllable can determine its meaning.

## Acknowledgments

This research was supported by the Deutsche Forschungsgemeinschaft (DFG), grant no. Ka 417/13-3. We want to thank Mrs. Juliane Grosser, who helped run the experiment, as well as Mrs. Anette Brechtel, for the preparation of the English version of this manuscript.

## References

- Burnett, T. A., Freedland, M. B., & Larson, C. R.** (1998). Voice F0 responses to manipulation in pitch feedback. *Journal of the Acoustical Society of America*, *103*, 3153–3161.
- Burnett, T. A., Senner, J. E., & Larson, C. R.** (1997). Voice F0 responses to pitch-shifted auditory feedback: A preliminary study. *Journal of Voice*, *11*(2), 202–211.
- Childers, D. G., & Krishnamurthy, A. K.** (1984). A critical review of electroglottography. *CRC Critical Reviews in Biomedical Engineering*, *12*(2), 131–161.
- Cowie, R., Douglas-Cowie, E., & Kerr, A. G.** (1982). A study of speech deterioration in post-lingually deafened adults. *Journal of Laryngology and Otology*, *96*(2), 101–112.
- Elman, J. L.** (1981). Effects of frequency-shifted feedback on the pitch of vocal productions. *Journal of the Acoustical Society of America*, *70*, 45–50.
- Houde, J. F., & Jordan, M. I.** (1998). Sensorimotor adaptation in speech production. *Science*, *279*, 1213–1216.
- Jäncke, L.** (1989). *Die Bedeutung der auditiven Rückmeldung des eigenen Sprechens für die Kontrolle der Phonationsdauer*. Inaugural-Dissertation, Heinrich-Heine-Universität Düsseldorf.
- Jäncke, L.** (1991). The “audio-phonatoric coupling” in stuttering and nonstuttering adults: Experimental contributions. In H. F. M. Peters, W. Hulstijn, & C. W. Starkweather (Eds.), *Speech motor control and stuttering* (pp. 171–180). Amsterdam: Elsevier Science.
- Kalveram, K. T.** (1984). Geschlechterunterschiede bei der audio-motorischen Kontrolle der Phonation. *Zeitschrift für experimentelle und angewandte Psychologie*, *31*(1), 39–47.
- Kalveram, K. T.** (1991). How pathological audio-phonatoric coupling induces stuttering. A model of speech flow control. In H. F. M. Peters, W. Hulstijn, & C. W. Starkweather (Eds.), *Speech motor control and stuttering* (pp. 163–170). Amsterdam: Elsevier Science.
- Kalveram, K. T., & Jäncke, L.** (1989). Vowel duration and voice onset time for stressed and nonstressed syllables in stutterers under delayed auditory feedback condition. *Folia Phoniatica*, *41*, 30–42.
- Kawahara, H.** (1994). Interactions between speech production and perception under auditory feedback perturbations on fundamental frequencies. *Journal of the Acoustical Society of Japan*, *15*, 201.
- Larson, C. R.** (1998). Cross-modality influences in speech motor control: The use of pitch shifting for the study of F0 control. *Journal of Communication Disorders*, *31*, 489–503.
- Natke, U.** (1999a). *Die audio-phonatorische Kopplung als grundlegender Mechanismus der Sprechflußkontrolle bei stotternden und nichtstotternden Personen*. Hannover: Unser Verl.
- Natke, U.** (1999b). Die Kontrolle der Phonationsdauer bei stotternden und nichtstotternden Personen: Einfluss der Rückmeldelautstärke und Adaptation (Control of phonation duration in stuttering and nonstuttering persons: Influence of feedback loudness and adaptation). *Sprache - Stimme - Gehör*, *23*, 198–205.
- Natke, U., & Kalveram, K. Th.** (2001). Fundamental frequency and vowel duration under frequency shifted auditory feedback in stuttering and nonstuttering adults. In H.-G. Bosshardt, J. S. Yaruss, & H. F. M. Peters (Eds.), *Fluency disorders: Theory, research, treatment and self-help. Proceedings of the Third World Congress on Fluency Disorders in Nyborg, Denmark* (pp. 66–71). Nijmegen: Nijmegen University Press.
- Parlitz, D., & Bangert, M.** (1999, June). *Short and medium motor responses to auditory pitch shift: Latency measurements of the professional musician's audio-motor loop for intonation*. Paper presented at the 16<sup>th</sup> International Congress on Acoustics and the 137<sup>th</sup> meeting of the Acoustical Society of America, Seattle, WA.
- Sapir, S., McClean, M. D., & Larson, C. R.** (1983). Human laryngeal responses to auditory stimulation. *Journal of the Acoustical Society of America*, *73*, 315–321.
- Siegel, G., & Pick, H. L.** (1974). Auditory feedback in the regulation of voice. *Journal of the Acoustical Society of America*, *56*, 1618–1624.
- Sundberg, J.** (1987). *The science of the singing voice*. Dekalb: Northern Illinois University Press.

Received August 9, 2000

Accepted February 8, 2001

DOI: 10.1044/1092-4399(2001/ )

Contact author: Dr. rer. nat. Ulrich Natke, Institute of Experimental Psychology, Section of Cybernetical Psychology and Psychobiology, Heinrich-Heine-University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany. Email: natke@uni-duesseldorf.de