AN INVESTIGATION OF A SPECIAL TYPE OF ACCENTUATION IN RIPUARIAN DIALECTS BY COMPUTER SIMULATION OF SPEECH PRODUCTION

B. J. Kröger, G. Heike, C. Opgen-Rhein, R. Greisbach and O. Esser

Institut für Phonetik der Universität zu Köln, Köln, Germany.

ABSTRACT

Recent studies on "Rheinische Schärfung", an accentuation feature in Ripuarian dialects, rely mainly on perceptual or acoustic analysis. Here an attempt is discussed to model this predominantly phonatory phenomenon by coupling a two-mass-model of the glottis to the Cologne articulatory speech synthesis system.

1. INTRODUCTION

The characteristic accentuation phenomenon of Ripuarian dialects, "Schärfung" ("sharpening"), is investigated by means of resynthesis. Hypotheses about the "sharpening" feature in the syllable [ɔ:s] ("carrion") were tested by synthesis in two ways: by glottal abduction and by glottal adduction.

2. ACOUSTIC CUES AND A PRODUCTION HYPOTHESIS

"sharpening" is an accentuation phenomenon whose phonetic features characterize additionally the nucleus of stressed syllables. It occurs either in long vowels or in diphthongs or in short vowels followed by a sonorant ([hy:] "height", [zɛi] "sieve", [ɔ:s] "carrion", [vo:t] "rage", [al] "all", [hyŋ]"dogs", [tant]"aunt")

The major acoustic cues of "sharpened" vs. "unsharpened" long vowels before voiceless consonants are [1]:

a) shortening of vowel duration;

b) a zero-intensity interval between vowel and following fricative;

c) a marked intensity decrease in the vowel segment;

d) a marked decrease in fundamental frequency in the vowel segment.

In terms of speech production "sharpening" is related to phonatory rather than to supraglottal articulatory activity [2]. There are in principle two maneuvers which lead to the acoustic features in question: a glottal abduction gesture and a glottal adduction gesture. They both begin in the middle of the vowel and last to its end.

3. THE SYNTHESIS MODEL

The articulatory speech synthesis model used was developed at the Institute of Phonetics, University of Cologne [3,4,5]. On the segmental level phonetic segments are put in. The model generates a set of articulatory control parameters (e.g. tongue height, tongue position, jaw opening) and three phonatory control parameters (lung pressure, cord tension, abduction area) which are continuous in time.

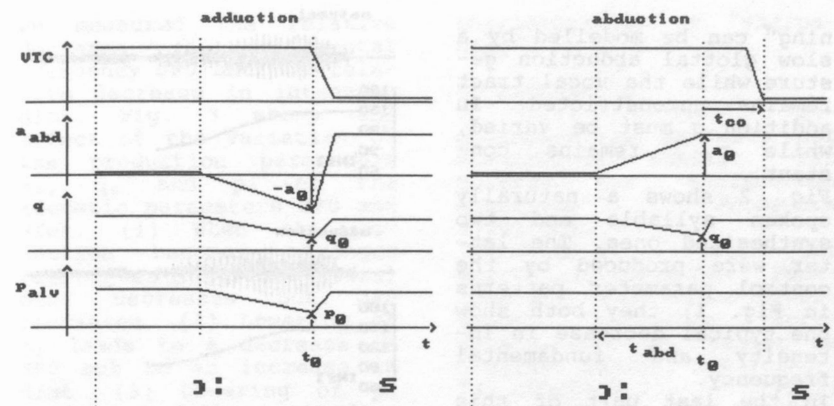

Figure 1: The control and production parameters for the "sharpening" for glottal abduction and glottal adduction.

The vocal tract is modelled with the Kelly-Lochbaum reflection-type line model; it simulates the wave propagation in the vocal tract by scattering partial waves at impedance discontinuities.

The self-oscillating glottis model is of major importance in this context. It comprises a static (nonoscillating) part or bypass and an oscillating part, the two-mass model [6]. While the two masses of the latter differentiate between the motion in the lower and the upper glottal region, the combination of the static plus the oscillating part differentiates between the motion in the posterior and anterior part. The bypass models the posterior portion, the less flexible part of the vocal folds. The control parameter abduction area represents the vocal fold position and changes the opening area of the bypass and the phonation neutral area of the two-mass model. Negative abduction area produces a glottal state in which the vocal folds are compressed medially.

4. "SHARPENING" IN THE PRODUCTION MODEL

Fig. 1 shows the pattern of the most important control parameters for [ɔ:s] for glottal abduction and glottal adduction. The vocal tract constriction VTC is given by the control parameter tongue tip height. The other control parameters are: abduction area a_{abd}, cord tension q, lung pressure (alveolar pressure) p_{alv}.

In the case of adduction the abduction area becomes negative at the end of the vowel, and medial compression is produced. At the same time, cord tension and lung pressure decrease, resulting in the decrease of F0 and signal intensity.

In the other case "sharpe-

ning" can be modelled by a slow glottal abduction gesture while the vocal tract remains unconstricted. In addition q must be varied, while p_{alv} remains constant.
Fig. 2 shows a naturally spoken syllable and two synthesized ones. The latter were produced by the control parameter patterns in Fig. 1; they both show the typical decrease in intensity and fundamental frequency.
In the last part of this study we investigated the influence of the most important production parameters on the acoustic cues of "sharpening".
In the case of glottal adduction these are the abduction area a_0, the cord tension q_0 and the lung pressure p_0 at instant t_0 (see Fig. 1). The time t_0 is defined as the instant at which the adduction/abduction gesture ends.

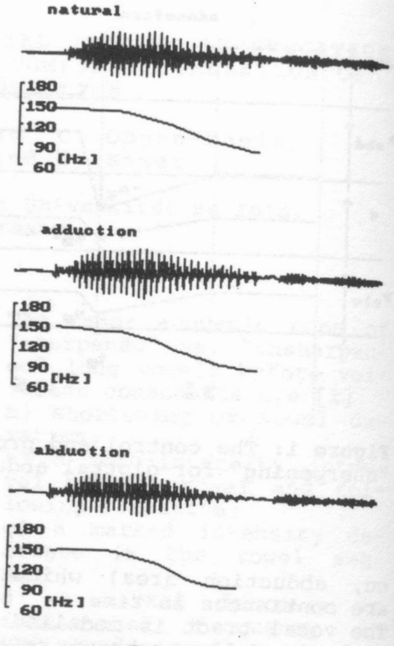

Figure 2: Resynthesis of "sharpening" in [ɔ:s] by glottal adduction and glottal abduction.

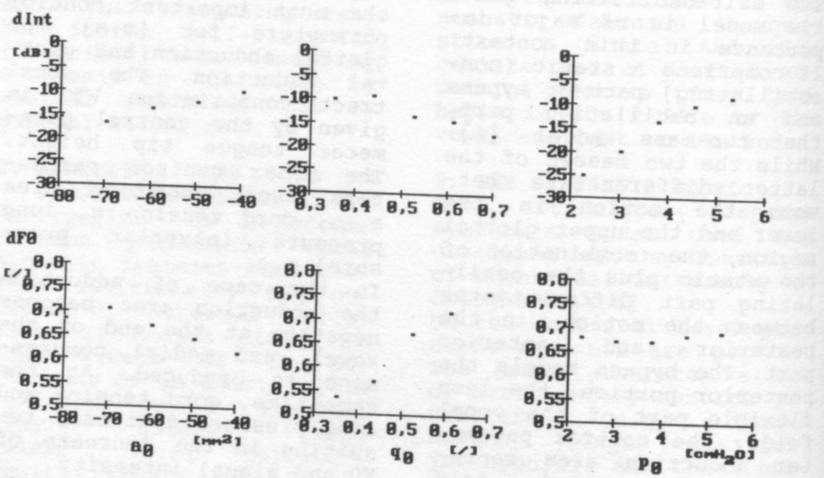

Figure 3: Acoustic parameters as function of production parameters. Case: adduction.

We measured the relative decrease in fundamental frequency dF0 and the relative decrease in intensity dInt. Fig. 3 shows the effect of the variation of the production parameters a_0, q_0 and p_0 on the acoustic parameters dF0 and dInt. (1) When the adduction becomes stronger (a_0 higher negative values) dInt decreases but dF0 increases. (2) Lowering of q_0 leads to a decrease in dF0 but to an increase in dInt. (3) Lowering of p_0 leads to a decrease in dInt while dF0 remains relatively unchanged.
So during strong adduction it is necessary to lower the cord tension and the lung pressure to get the acoustic features of the "sharpening".

5. DISCUSSION

Resynthesis of "sharpenend" syllables by glottal abduction or adduction was done. In the latter case, mechanical compression (a_{abd} < 0), low cord tension and a decrease in lung pressure are necessary. Similar physiological features were found for the Danish "stød" [7]. The production mechanisms are those of "creaky voice" phonation. Modelling "sharpening" by glottal abduction seems to be the less promising way since at the end of the vowel there is a relatively high air flow through glottis and vocal tract which may produce aspiration. But both production mechanisms generate the main acoustic cues characteristic of "sharpening".

6. REFERENCES

[1] HEIKE,G. (1962), "Suprasegmentale Merkmale der Stadtkölner Mundart. Ein Beitrag zur 'Rheinischen Schärfung'", *Phonetica 8*,147-165.
[2] HEIKE,G. (1988), *"Zur wortunterscheidenden Funktion der rheinischen Schärfung"*, Deutscher Wortschatz, 677-686, Walter de Gruyter: Berlin, New York.
[3] KRÖGER,B.J. (1989), *"Die Synthese der weiblichen Stimme unter besonderer Berücksichtigung der Phonation"*, Dissertation, Köln.
[4] HEIKE,G., GREISBACH,R., HILGER,S., KRÖGER,B.J. (1989), *"Speech synthesis by acoustic control"*, Eurospeech 89, Vol.2, 20-22, European Conference on Speech Communication and Technoloy, Paris.
[5] KRÖGER,B.J., OPGEN-RHEIN,C. (1990), "Das physiologisch akustische Modell AKUSYN des artikulatorischen Sprachsynthesesystems ASSCO", *IPKöln-Berichte 16*, 69-117.
[6] ISHIZAKA, K., FLANAGAN, J.L. (1972), "Synthesis of voiced sounds from a two-mass model of the vocal cords", *Bell System Technical Journal 51*, 1233-1268.
[7] FISCHER-JØRGENSEN, E. (1989), "Phonetic analysis of the Stød in Standard Danish", *Phonetica 46*, 1-59.