PERCEPTION OF VOWEL QUANTITY IN DYSARTHRIC SPEECH - INTERACTION OF ACOUSTIC, ARTICULATORY AND PERCEPTIVE TIMING MECHANISMS

Gabriele Scharf
Edith-Stein-Rehabilitation Clinic, Bad Bergzabern, Germany

ABSTRACT

It is the purpose of the present study to investigate the influence of different acoustic and articulatory timing mechanisms in dysarthric speech on the perceivability of the phonological vowel quantity contrast in German. Sentence utterances of patients with dysarthria due to cerebellar atrophy and Parkinson's disease were auditively judged as well as analyzed on the acoustic and articulatory level. Data analysis indicated an interaction of articulatory timing (positioning of vowel onset in the labial opening gesture), acoustic timing (vowel duration and speech rate) and perceptual ratings of vowel quantity. The results suggested that different perceptive strategies were applied dependent on the type of dysarthria. Investigation of the interrelation between segmental and suprasegmental speech timing in dysarthria on the one hand and auditive perception performance on the other hand might shed light on general mechanisms in the perception of segmental contrasts.

1. INTRODUCTION

Even though the assessment of dysarthria traditionally bases on auditive judgements [4], very little is known about the specific processes running in the perceptive system of listeners of dysarthric speech. As demonstrated by numerous studies, dysarthric speech production is characterized by deviant articulatory patterns as well as by specific segmental and suprasegmental deviations of the acoustic signal [6]. But as far as I know no experimental study so far systematically investigated the influence of these pathological speech characteristics on the perception of dysarthrics' utterances.

Deviant vowel duration as well as distortion of vowel formant structure are regarded as typical features of different dysarthia types [4]. As could be shown by [1,2] and others, cerebellar dysarthria is characterized by overall slowed speech rate. In contrast, hypokinetic dysarthria in Parkinson's disease is often related to increased speech rate [3]. Data of [8] indicated that the coordination of vowel onset and articulatory gesture may be impaired in dysarthric patients as well.

In the present study the influence of speech characteristics of cerebellar and Parkinsonian dysarthria on the perceivability of the phonological vowel quantity contrast in German schall be investigated. Possible effects of vowel duration, vowel quality and speech rate were of particular interest. In addition, timing of the vowel onset in the articulatory cycle was examined.

2. METHODS

2.1. Subjects

10 dysarthric patients with idiopathic Parkinson's disease (PARK), 11 patients with dysarthria due to cerebellar ataxia (CERE) and 15 normal controls without any history of neurological disease or speech disturbance (NORM) produced test utterances in the present study.

2.2. Material

Test sentences were of the form /*ich habe gepVpe gelesen*/ (I've read gepVpe). In a first experiment the target vowel *V* was varied with regard to vowel quantity (tense/lax) and vowel type ("a", "i", "u") resulting in 6 different target vowels: [a:, a, i:, I, u:, U]. In experiment 2 vowel quantity was varied in only one vowel type [a:, a]. Additionaly, the subjects were instructed to realize the test sentences blockwise in three different rates (as fast as possible, normal, very slow). Each experiment comprised 48 sentences per speaker out of 8 repetitions of each of the 6 conditions.

2.3. Acoustic analysis

The sentence productions were recorded on a DAT tape. Acoustic analysis was carried out by Computerized Speech Lab (CSL 4300, Kay Elemetrics / USA) and by Xwaves software on a Unix-Workinstation (Silicon Graphics), respectively (sampling rate 16 kHz, 16 bit). The following acoustic measures were considered for analysis: The durational vowel quantity contrast was determined by the mean quotient of the target vowel duration (VDUR) of tense and lax vowels (REL= VDUR tense / VDUR lax). As a quantity independent measure of speech rate the utterance duration minus duration of the syllable containing the target vowel was calculated (RATE). The adaptation of vowel duration and speech rate is reflected by the parameter (VOKREL = VDUR / RATE). Formant frequencies (F1, F2, F3) were registered at the midpoint of the target vowel (BURG-algorithm, Hamming-window, 20 coefficients). The quotient FQ = (F2 / F1) was calculated as a measure of vowel quality for tense and lax vowels separately. FREL was defined as quotient of (FQ tense / FQ lax) and was considered to be a measure of the vowel quality contrast between tense and lax vowel counterparts.

2.4. Kinematic analysis

An optoelectronic movement analysis system (ELITE, Milano/Italy) with two infrared cameras was used to track lip movements at a sampling frequency of 100 Hz. Reflective markers were attached to the upper and lower lips. Figure 1 shows the acoustic signal and the labial trajectory (Dislip) of the target sequence /pap/. The relative position of the vowel onset in the labial opening gesture was examined by the formula (VOKPOS = OPVOK / OPDUR), that is the quotient of the time difference between onset of the opening gesture and vowel onset (OPVOK) and the duration of the opening gesture (OPDUR). Since lip opening is the crucial articulatory feature solely in the production of vowel "a", VOKPOS is only calculated for this vowel type.

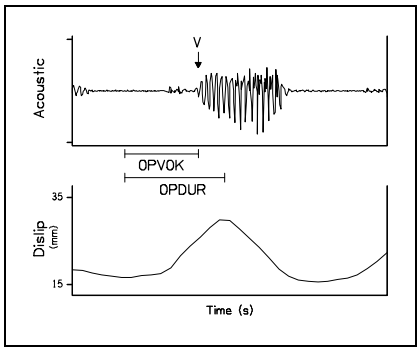

Figure1. Acoustic signal and labial kinematics of the target sequence /pap/

2.5. Perceptual rating

Two female speech-language pathologists (26 and 28 years old) who didn't report subjectively a hearing disorder listened to the sentence utterances binaurally presented on a headphone at comfortable listening level. Perceptual error scores (ERR) were calculated from the auditive rating of vowel quantity of the two listeners and were related to the different acoustic and articulatory measures. Interrelation between ERR and the different acoustic and articulatory measures was investigated by pairwise correlation tests.

3. RESULTS

Errors in the perception of vowel quantity in NORM occurred only occasionally. Perceptual ratings of vowel quantity in the utterances of both dysarthric groups were characterized by substantially higher error rates than in NORM (CERE and PARK: $p<0.0001$). This is the case for both experiments. The errors were heterogeneously distributed to the different vowel types and speech rate conditions (Fig. 2). Table 1 provides an overview of the results of the correlation tests.

Table 1. Pearson coefficient and Bonferroni probabilities for the correlation of perceptual error rates with the different acoustic and articulatory parameters (*: $p<0.05$, **: $p<0.01$).

	Vowel	NORM	CERE	PARK
REL	**„a“**	n.s.	* -0.667	** -0.865
	„i“	* -0.602	n.s.	* 0.712
	„u“	n.s.	n.s.	* -0.714
FREL	**„a“**	n.s.	n.s.	n.s.
	„i“	n.s.	n.s.	n.s.
	„u“	n.s.	n.s.	** -0.879
VOKPOS	**„a“**	n.s.	n.s.	** 0.869

3.1. Experiment 1

3.1.1. Vowel duration Normal speakers realized a relatively stable durational contrast between tense and lax vowels (REL) of about 2:1 for the three vowel types. Both dysarthric groups showed a reduced durational quotient in all vowel types (CERE: $p<0.05$, PARK: $p<0.01$). A relation of low REL values to perceptual errors in NORM were found only for vowel "i". CERE showed a correlation between error rate and reduced REL in vowel "a". In PARK a negative correlation betwen durational vowel contrast and error frequency was found for all vowel types.

3.1.2. Vowel quality No significant differences between groups could be found with respect to the formant frequencies F1, F2 or F3 of the target vowels. Apart from a negative correlation between ERR and FREL for vowel "u" in PARK (Table 1) there was no significant correlation of perceptual error rate and this measure of vowel quality contrast in one of the groups.

3.1.3. Position of vowel onset in labial opening The position of the vowel onset in the labial opening gesture (VOKPOS) proved to be a stable acoustic-articulatory correlate of the vowel quantity contrast - at least for vowel type "a" - in NORM: All normal speakers without exception positioned the vowel onset of lax vowel [a] remarkably earlier in the labial opening gesture than in the case of the tense vowel [a:] [8]. Both groups of dysarthric speakers showed a reduced vowel quantity contrast with respect to this parameter. But only in the case of PARK a positive correlation could be found between VOKPOS and perceptual error rate (0.869, $p<0.01$).

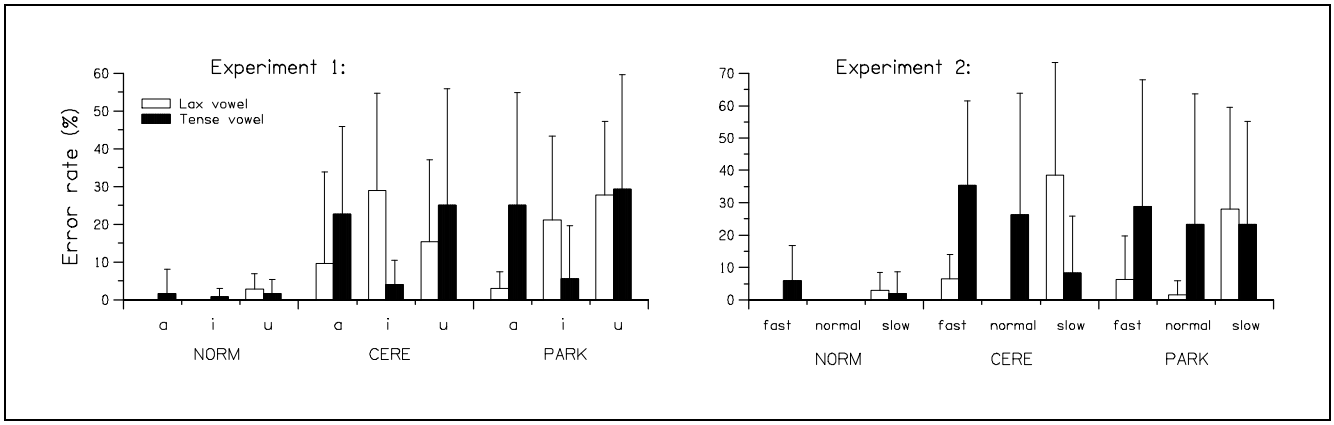

Figure 2. Frequency of perceptual errors of vowel quantity (%) in experiment 1 and 2

3.2. Experiment 2

In the analysis of experiment 2 data of the three speech rate conditions (fast, normal, slow) was pooled for each speaker.

3.2.1. Vowel duration and speech rate Figure 3 illustrates the differential adaptation of vowel duration to speech rate variation for tense and lax vowels in NORM (top). Vowel duration increased with reduced speech rate - the more for tense vowels in comparison to the lax vowels. This is indicated by the diverging regression lines for the two vowel classes. As indicated by the nearly parallel lines in CERE (mid) and the almost converging lines in PARK (bottom) these two patient groups were not able to sufficiently adapt vowel duration in dependence on speech rate. This lack of durational adaptation has to be interpreted as timing deficit of the dysarthrics. An influence of this mistiming on the perception of the vowel quantity contrast is proved by the correlation coefficients between error rates and VOKREL (Table1): In PARK a correlation between ERR and VOKREL was found for both vowel classes (tense and lax) whereas in CERE the high error rate correlated only with the lacking adaptation of durations of tense vowels.

Table 1. Pearson correlation and Bonferroni probabilities for the correlation of perceptual error rate with VOKREL (definition see text) (*: p<0.05, **: p<0.01).

	NORM		CERE		PARK	
	[a:]	[a]	[a:]	[a]	[a:]	[a]
VOKREL	n.s.	n.s.	*	n.s.	*	*
			-0.546		-0.538	-0.496

3.3. Summary

Substantially higher error rates in the perception of vowel quantity were observed for both groups of dysarthric speakers. Errors were heterogeneously distributed to the different experimental conditions (vowel quantities, vowel types, speech rates). The present data cannot fully explane the complex error pattern. However, some of the results indicate

(1) Increased error rates in the perception of vowel quantity could be related to a reduced durational vowel quantity contrast in PARK for "a" and "i", in CERE only for "a".

(2) Deviant formant frequencies (like reduced vowel quality contrast or general vowel centralization) couldn't account for the generally high perceptual error scores in the dysarthrics. (PARK showed the only exception with respect to the „u“ vowels.) In contrast, timing deficits seemed to be more salient.

(3) In PARK the perceptual error pattern of "a" was correlated with the deviant pattern of VOKPOS, the position of vowel onset in the labial opening gesture. This was not the case for CERE.

(4) In PARK the perception of tense *and* lax vowels seemed to be influenced by the insufficient adaptation of target vowel duration and actual speech rate (VOKREL). In CERE an influence of this kind is only suggested in tense vowels.

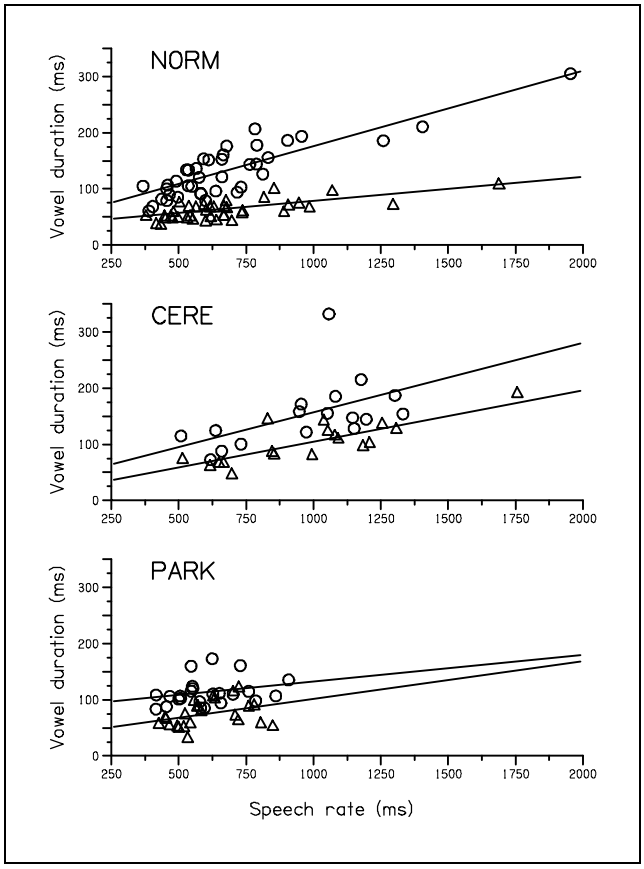

Figure 3. Target vowel duration and speech rate (defined as utterance duration minus target syllable duration). 3 symbols for each speaker (fast, normal, slow condition). Circles: tense vowels, triangles: lax vowels.

4. CONCLUSIONS

The perception of the segmental phonological contrast of vowel quantity in German is dependent on complex segmental and suprasegmental timing mechanisms on articulatory and acoustic level. Impairment of these timing processes characterizes the speech of dysarthric patients with cerebellar ataxia and Parkinson`s disease, respectively, and might be responsible for the reduced perceivability of the quantity contrast in their utterances. The durational quotient of tense and lax vowels (REL) was of crucial importance. The exact realization of a durational vowel contrast requires exact timing of acoustic and articulatory events [5]. It shall be argued that the position of the vowel onset in the labial opening gesture, at least for vowel "a", represents a crucial categorical feature of vowel quantity [8]: The synchronization of vowel onset and maximum lip opening in lax vowels - uniformly shown by all normal speakers - might be interpreted as the best articulatory configuration for the production of maximum sonority. It is conceivable that reaching or not reaching of this articulatory target might be of perceptual relevance. Pathological deviation of this pattern may lead to higher perceptual error rates as demonstrated by PARK. However, no interrelation between the deviant VOKPOS patterns and the perceptual scores was shown in CERE. Diverging results of CERE and PARK might indicate that different perceptual strategies are used for different types of dysarthria. Further perception studies on dysarthric speech are needed in order to test this hypothesis. If it could be confirmed, this would be of great importance for future management of dysarthria.

An important suprasegmental timing mechanism in the realization of vowel quantity is represented by the differential tuning of vowel duration and actual speech rate. This suprasegmental timing process is specifically disturbed in dysarthria: Timing demands of the segmental contrast of vowel quantity interfere with the demands of speech rate. Correlation results suggest that on the part of the listeners of dysarthrics' utterances corresponding rate normalization takes place [7] resulting in frequent misinterpretations. It has to be emphasized that no such correlation was found in the listeners of normal speakers. Therefore, one might assume that in the perception of normal speech other processes work than in the perception of pathological speech. It can be speculated that the hearer of a speech utterance is able to choose between different perceptual strategies.

REFERENCES

[1] Ackermann, H. and Hertrich, I. 1994. Speech rate and rhythm in cerebellar dysarthria: An acoustic analysis of syllabic timing. *Folia Phoniatrica et Logopaedica* 46, 70-78.

[2] Ackermann, H., Hertrich, I. and Scharf, G. 1995. Kinematic analysis of lower lip movements in ataxia dysarthria. *Journal of Speech and Hearing Research* 38, 1252-1259.

[3] Ackermann, H. and Ziegler, W. 1991. Articulatory deficits in Parkinsonian dysarthria: an acoustic analysis. *Journal of Neurology, Neurosurgery, and Psychiatry* 54, 1093-1098.

[4] Darley, F.L., Aronson, A.E. and Brown, J.R. 1975. *Motor Speech Disorders*. Philadelphia: Saunders.

[5] Hertrich, I. and Ackermann, H. 1997. Articulatory control of phonological vowel length contrasts: Kinematic analysis of labial gestures. *Journal of the Acoustical Society of America* 102, 523-536.

[6] Kent, R.D. 1990. The acoustic and physiologic characteristics of neurologically impaired speech movements. In Hardcastle, W.J. & Marchal, A. (eds.), *Speech Production and Speech Modelling*. Dordrecht: Kluwer Academic Puclishing, vol.55.

[7] Pind, J. 1995. Speaking rate, voice-onset time and quantity: The search for higher-order invariants for two Icelandic speech cues. *Perception & Psychophysics* 57, 291-304.

[8] Scharf, G. 1999. (forthcoming) Timing-Defizit bei Dysarthrophonie. Die akustische und artikulatorische Realisierung von Vokalquantität bei Patienten mit Morbus Parkinson und zerebellärer Ataxie. *Arbeitspapiere des Instituts für Maschinelle Sprachverarbeitung AIMS. University of Stuttgart.*