DURATIONAL PATTERN OF RUSSIAN SYNTAGMA: THE STANDARD SCHEME AND ITS MODIFICATIONS

Olga Krivnova

Department of Philology, Moscow State University
Moscow, USSR, 119899

ABSTRACT

In the present paper it is suggested the idea of the existence of a pronunciational background related to rhythmical speech organization. In temporal aspect this background is realized in particular through some standard scheme of phonosyntagma pattern. On the basis of coherent Russian text the most important qualitative features of the standard scheme are revealed and a numerical model and statistical characteristics of its realization are presented.

INTRODUCTION

The principle of speech rhythmical organization suggests forming a certain pronunciational background in speech signal. Regular reproduction of the elements realizing this background leads to disintegration of speech into various phonetic constituents /phonowords, syntagmas, phrases, periods/ with their specific inner structure periodically recurring in rhythmically organized speech. Rhythm-forming elements having various physical realization attach important functions to acoustic speech parameters - functions of form construction and integrity of speech units. The subject functions are intrinsic to duration as well. In this paper a possibility of presentation of phonosyntagma durational patterns in Russian as a result of realization of a standard scheme and its regular modifications is considered. Besides, the standard scheme is viewed as temporal stereotype which, being a part of pronunciational background and skills of a native speaker, plays an important role in creating phonetic integrity of the syntagma.

STANDARD SCHEME OF SYNTAGMA TEMPORAL ORGANIZATION

Qualitative description

A great number of phonetic experimental research as well as speech synthesis practice demonstrate that the general tendency of syntagma temporal arrangement in Russian is related to forming a positional contrast or, in other words, differences in duration of various components of a phonetic word depending on its position in the syntagma. Various language material shows an almost universal character of the positional factor when analysing its influence on word durational characteristics within syntagma. The positional contrast in a Russian syntagma is clearly detected not only when isolated phrases are pronounced but in running speech as well. Thus we can speak about the existence of a corresponding temporal stereotype or a standard realization scheme of the positional contrast. At the same time up to now the positional contrast has been studied on separate phrases with a limited set of rhythmical word-patterns and sound composition. But if we try to analyse running speech, the data obtained from such a material are not sufficient. In view of the abovementioned facts the first part of our research was devoted to the qualitative analysis of intrasyntagmatic duration relations in coherent text. Our purpose was first of all to enrich and make more accurate the experimental data obtained earlier.

The study was carried out on a corpus of syntagmas singled out as a result of auditory analysis of coherent scientific Russian text read by an announcer /standard Moscow speaker/ with moderate individual speaking rate.

Syntagmas of various length with a main stressed word in final position were selected for this analysis. Total volume of the sample comprised 438 units. Sound durations were measured according to oscillograms /registered at film speed of 100 mm/sec/. In order to make the segmentation procedure easier speech signal and intensity curve were recorded on the film simultaneously.

As a result of the analysis the following peculiarities of the positional contrast /PC/ were revealed:

1. A normative /statistically predominant/ way of a PC realization is time shortening of a word in the syntagma non-final position. The shortening is achieved by fast speaking rate of these words while maintaining normal speaking rate for words in the final position and single-word syntagmas. The following symbols can be used to designate the way of the PC construction: F /nf/ - N /f/. Further we'll speak about it as of standard temporal scheme /STS/ of the syntagma. Normative nature of the subject scheme is demonstrated by the results of the comparison of the word speaking rate characteristics in various syntagma positions with data on individual speaking rate differences in Russian. Tables I and 2 present data on average sound duration for various individual speaking rates and for words in different syntagma positions accordingly.

Table I. Average sound duration /in msec/ of individual speaking rate in Russian /I/

INDIVIDUAL SPEAKING RATE	FAST	NORMAL	SLOW
GENERAL	65	65-73	73
VARIATION WITHIN NORMAL GENERAL TEMPO	60,4-1,4	74,2-3,0	87,1-2,2

Table 2. Average sound duration /in msec/ for words in different syntagma positions

NUMBER OF WORDS IN A SYNTAGMA	POSITION INITIAL	MEDIAL	FINAL
2	60,0-2,0	---	77,0-3,0
3	59,0-1,4	58,0-1,2	76,0-2,0

Data on word stress perception in a Russian syntagma /2/ make it possible to assume that the scheme F /nf/ - N /f/ being a standard method of the PC realization takes part in the formation of a syntagma accent pattern /component known as syntagmatic stress/. In this connection one may notice that in works on Russian phonetics /3/, /4/ the idea that an increase of the speaking rate can be considered as a source of numerical and qualitative reduction of vowels in a word was put forward more than once. The notion that the fast tempo of pronunciation of syntagma non-final words decreases their prominence level and creates the contrast needed for syntagmatic stress is a natural widening of this idea.

2. Apart from the positional factor, the word duration depends also on such shortening factors as the number of syllables per word and the word distance from the syntagma beginning point. Under the total influence of all shortening factors duration is decreasing in a nonlinear manner showing what is known as "an incompressibility effect" /5/. This effect is clearly seen in syntagma non-final positions /fig. I/.

3. Minimal vowel durations characterizing the incompressibility effect are close to certain temporal perception constants. For example, the minimal duration of an unstressed vowel /$T_{min} \approx$ 30-40 msec/ is close to the threshold value of its detection under any consonant environment /6/. The minimal duration of a stressed vowel /$T_{min} \approx$ 70-90 msec/ is close to threshold value needed for its phoneme running identification /6/. It follows that vowel duration in syntagma non-final words is constrained within limitations which, first of all, provide a possibility of correct identification of rhythmic type of the word and recognition of its stressed vowel. It is also worth mentioning that minimal durations of stressed and unstressed vowels relate as 2 to I.

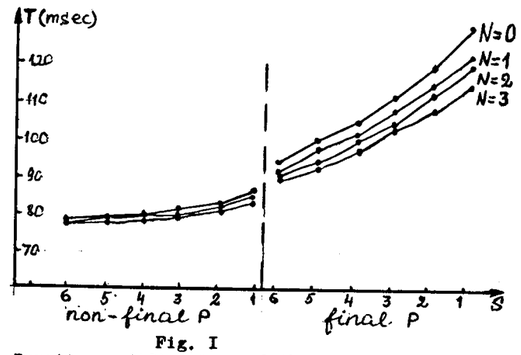

Fig. I
Duration modification of the stressed vowel in syntagma under the influence of the following factors: position /P/, number of syllables per word /S/, word distance from the syntagma beginning point measured in lexical stresses preceded /N/. Average values disregarding vowel quality distinctions.

4. Duration boundary values which separate stressed vowel realizations in non-final and final words /$T_{and} \approx$90-100 msec/ are close to phoneme boundary values obtained for languages with phonemic contrast in vowel length /7/, /8/. This makes it possible to speak about different duration categories in Russian speech as well. The analysis shows that a vowel of a mid length is realized when it is stressed, belongs to syntagma final word but not followed by a pause. In that case its duration is about 3 times longer than the critical value of vowel detection. If the latter is classified to be supershort, then vowel realizations of other categories are characterized by the following relations of durations - I/supershort/: 2/short/: 3/mid/: 4/long/.

It is obvious that STS supposes the stressed vowels of non-final words in syntagma to be short and the stressed vowels of final words to be mid.

5. Positional shortening of non-final words in Russian syntagma is of an asymmetric nature: the shortening is mostly noticeable in vowels in the word terminal part beginning with its stressed vowel /"swallowing" of word terminals is an extreme manifestation of this peculiarity/.

Asymmetry leads to smoothing and actually to the loss of temporal contrast of prestressed and stressed vowels in the syntagma non-final word and this apparently hampers correct stress identification.

Numerical model

The second part of our study is devoted to the development of a numerical model of the PC standard temporal scheme. Such a model is of great interest from the various points of view. Generality of the positional factor in its influence on speech temporal characteristics causes us to think that specific linguistic features of the PC are imbodied in its numerical parameters.

The results of the preliminary qualitative

analysis give us an idea of the general form of the STS numerical model. While determining concrete numerical estimates we limited ourselves to the consideration of stressed vowels /V́/. Data of their duration were obtained by using the same corpus of syntagmas that has already been described above.

The proposed STS model accounts for all the shortening factors: word position in the syntagma /P/, the number of syllables per word /S/, the number of preceding lexical stresses /N/.

Elongation factors such as absolute prepausal position and position under semantic accent are not taken into account. Consonant environment that influences V́ is not taken into consideration as well. Thus, formulae /I, 2/ give mean duration to the non-prepausal V́ in a syntagma with final main stress when there is no pronounced semantic accent in it.

The general form of the model was chosen on the basis of the notions that were used before in some generative models of speech temporal organization /9/, /I0/.

These notions are as follows:

I. For every stressed vowel of a given phonetic quality there can be singled out two specific realizations: first in the context where both elongation and shortening factors are absent and second in the context where shortening factors have their maximum effect. In the first case it is natural to regard the vowel duration as its intrinsic duration $/T_o/$, and in the second case as its critical duration $/T_{min}/$. The difference $/T_o - T_{min}/$ which characterizes a stressed vowel temporal potential can be called its residual duration.

2. The shortening factors P, S, N effect $/T_o - T_{min}/$ and besides they act independently and irrespective of vowel quality.

3. The factors S and N shorten the stressed vowel recursively, i.e. their shortening effect is realized cyclicly as a function of the corresponding variables.

In accordance with the accepted assumptions the V́ duration in the absence of elongation factors can be presented as follows:

$$T_{S,P,N}(\acute{V}) = (T_o - T_{min}) \cdot \alpha^{S-1} \cdot \beta^{P} \cdot \gamma^{N} + T_{min} \quad (1),$$

where S, N as defined above, P can acquire 2 meanings: 0 - for a word in syntagma final position and I - for a word in non-final position, α, β, γ - shortening coefficients.

As a result of the examination of V́ measured durations the α, β, γ -coefficients were approximated as 0,82; 0,20; 0,90 accordingly. In compliance with the obtained estimates the subject suprasegmental factors can be ordered according to their increasing shortening effect as follows: N, S and P. Such an order corresponds to the data presented in other publications.

With the received estimates, the STS model for syntagma stressed vowels looks like this:

$$T_{S,P,N}(\acute{V}) = (T_o - T_{min}) \cdot 0{,}82^{S-1} \cdot 0{,}20^{P} \cdot 0{,}90^{N} + T_{min} \quad (2).$$

In formula 2 one can find the incompressibility effect detected empirically - vowel reduction decreases gradually when the shortening effect is increased.

Identity of formula 2 was verified on a corpus of 505 syntagmas /the same speaker performed the reading/. As a deviation value of measured and calculated durations the following value was selected: $|T^{msd.}_{S,P,N}(\acute{V}) - T^{calc.}_{S,P,N}(\acute{V})|$. This value was defined on different generalization levels of durations measured. We have established 3 levels: I - a level of average values of a generalized V́ /disregarding its phonetic quality/: T_o=I30 msec, T_{min}=75 msec, II - a level of mean values of V́ with regard of close-open distinctions: T^{a}_{o}=I47 msec, T^{a}_{min}=85 msec; $T^{c,e}_{o}$=I34 msec, $T^{c,e}_{min}$=78 msec; $T^{i,u,e}_{o}$=II4 msec, $T^{i,u,e}_{min}$=66 msec, III - a level of V́ specific realizations: the values T_o and T_{min} are the same as in case II. Sample mean value $|T_{msd.} - T_{calc}|$, deviation and 95% confidence interval for the sample mean ($\pm 2\frac{\sigma}{\sqrt{n}}$) were calculated for each generalization level. Table 3 shows the calculation results.

Table 3. Statistical characteristics of $|T_{msd.} - T_{calc}|$ for different generalization levels of empirical data

STATISTICAL CHARACTERISTICS	GENERALIZATION LEVELS I	II	III
THE SAMPLE MEAN /msec/	5,6	7,4	I5,8
MEAN-SQUARE DEVIATION σ /msec/	7,2	7,2	I3,I
$\pm 2\frac{\sigma}{\sqrt{n}}$	2,2	I,4	0,8
n - NUMBER OF THE ITEMS	4I	96	I03I

Accounting that the accuracy of duration measurement didn't exceed 5 msec the agreement between the calculated and empirical data can be considered as good. Cases of essential deviations demand a close analysis: they give evidence of additional factors the effects of which are not taken into consideration in the proposed model.

STS realization in running speech

The syntagma STS F /nf/ - N /f/ is a component of a pronunciational background stipulated by speech rhythmical organization. It is obvious that a concrete realization of the scheme in a coherent text depends on various contextual conditions. Formula 2 /when used to calculate vowel temporal scales/ shows variation of absolute values of a stressed vowel duration in syntagmas that conform to the STS but at the same time have different verbal filling. It is also important to detect cases of STS disagreement and to reveal their sources. It is worth remembering that violation of general rhythm-forming tendencies is the most relevant informative and descriptive mean in coherent text. Thus the task of detecting the syntagmas corresponding to the STS and those having deviations from it arises.

In this research the same material which served as the basis for the STS numerical model construction was used for the elaboration of the given task. Analysis-by-synthesis procedure was utilized to solve the problem of whether a syntagma chosen arbitraly from the text corresponds to the STS. The latter was considered to be realized in the syntagma if its stressed vowels had the durations compatible with the values calculated by formula 2. Thus empirical data were interpreted taking into account specific phonetic conditions /vowel quality, word rhythmical type, syntagma length/ under which STS is realized. If there was incompatibility of calculated and empirical V́ durations in the syntagma we considered that there was an STS violation in it.

The most difficult problem which arose in the process of implementation of above-mentioned procedure is the definition of critical value sufficient to acknowledge essential divergence of empirical and calculated durations. The difficulty of the problem lies in the fact that very little is yet known about how a listener perceives, estimates and interprets duration differences. We decided to consider a divergence unessential if the value $\frac{|T_{msd} - T_{calc}|}{T_{calc}}$ didn't exceed 20%. It is worth noticing that despite some divergences the overall data at present show that the critical value of just-noticeable difference for the perceived change in vowel duration is about I5%-20% /9/.

Before presenting the results of the experimental research conducted it is worth showing theoretically possible types of deviations from the STS. Since the STS presupposes the realization of the stressed vowel in the syntagma nonfinal and final words in short and mid lengths accordingly, it is easy to see that at the level of stressed vowels the following deviations from the standard are possible:

A - absence of V́ reduction in the syntagma nonfinal word /i.e. realization of the mid or long category/,

B - V́ reduction in the syntagma final word /i.e. realization of the short category/,

C - V́ lengthening in the syntagma final word /i.e. realization of the long category/.

Table 5 displays statistical data /%/ on STS realization obtained on the basis of our research material.

Let's look at the section I first. We can see that when the main stressed word is in final position the STS is realized without any deviations in approximately half of the cases. Within the set of syntagmas with STS violations those with one deviation prevail, syntagmas with more than one deviation constitute only II% of the overall set. It follows that in syntagmas of this type the observed deviations are local modifications of the standard and are not the result of its modification as a whole. The same conclusions may be deduced for syntagmas settled in section II.

Table 5. STS realization in coherent text
I - syntagmas with main stressed word in final position /438 units/, II - syntagmas with main stressed word in non-final position /67 units/.

	NUMBER AND TYPES OF DEVIATIONS FROM THE STS 0	I			2			3
	0	A	B	C	A,A	A,B	A,C	A,C
I	45,9	9,6	I9,4	I4,8	0,9	6,6	2,3	0,5
II	28,4	I9,4	I9,4	3,0	3,0	I4,9	9,0	3,0

On the basis of literary data and the results of our research the following sources and positions of probable STS disagreements can be distinguished:

I. Absolute prepausal V́ position/for final word/.
2. Type of pitch contour /for word in main stressed position/.
3. Weakening or strengthening of a word prominence level /for word in any position/.
4. Syntagma position in regard to utterance external and internal boundaries /for word in final position and position of main stress/.

REFERENCES

/I/ Агафонова Л.С., Бондарко Л.В. и др. О некоторых характеристиках русской речи в зависимости от разных темпов произнесения. - В кн.: Слух и речь в норме и патологии. Л., I974, I, с. I0-I7.

/2/ Светозарова Н.Д. Интонационная система русского языка. Л., I982.

/3/ Чистович Л.А. и др. Речь. Артикуляция и восприятие. М.-Л., I965.

/4/ Бондарко Л.В. Фонетическое описание языка и фонологическое описание речи. Л., I98I.

/5/ Klatt D.H. Interaction between two factors that influence vowel duration. JASA, I983, v.54, p. II02-II04.

/6/ Чистович Л.А. и др. Восприятие речи человеком. Л., I976.

/7/ Fujisaki H., Nakamura K., Imoto T. Auditory perception of duration of speech and non-speech stimuli. Ann. Bull. of Res. Inst. of LPhUT, I973, 7, p. 45-64.

/8/ Nooteboom S.G. Contextual variation and perception of phonemic vowel length. - Preprints of the Sp. Com. Sem., Stokholm, I974, 3, p. I49-3.

/9/ Klatt D.H. Linguistic uses of segmental duration in English: acoustic and perceptual evidence. JASA, I976, v. 59, p. I208-2I.

/I0/ Lindblom B., Lyberg B., Holmgren K. Durational patterns of Swedish phonology: do they reflect short-term memory processes? Indiana Univ. Ling. Club, I98I.