INTONATIONAL PHRASING AND ITS ROLE IN SPEECH COMMUNICATION

Olga Krivnova

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

ABSTRACT

The theoretical claim of this study is that intonational phrasing /IP/ functions in speech communication as one of language orientation means.

The paper presents the results of a psycholinguistic experiment, whose aim was to examine sense segmentation devices in two different situations: I/ in a written text having no formal signs of segmentation, whatsoever, 2/ in an oral text based on normally organized written speech sample.

INTRODUCTION

The cognitive background of modern linguistics presupposes the view that language as a means of communication can be adequately analysed only with references to the general principles of human speech behaviour.

It should be noted in this connection that the treatment of IP /i.e. division of an utterance stretch into phonosyntagmas, phrases and more lengthy speech units/ in Russian and Soviet linguistics emphasizes the direct relationship between this phenomenon and human speaking activity, viewing the former as inseparable from speech production and comprehension process.

This view is based on the works of the prominent Russian scholar - academician L.V.Scherba, who defined the minimal IP unit - phonosyntagma - as a phonetic unit conveying a semantic whole in the speaking-thinking process. Recent studies on IP are aimed at providing further evidence for the general idea of IP's relationship with speech activity. It seems that problems concerned with IP can't be solved without investigating the general principles of processing and verbalization of conceptual information as they carried out by the speaker. Therefore we think it useful to consider some ideas emerging in the framework of current theories of speech acts.

One of the most widely-known attempts to formulate the general rules governing speech behaviour belongs to the American scientist H.P. Grice /I/. He advanced an idea that speech communication is based on the Cooperative Principle in accordance with which participants of the communicative process interested in its successful realization, perform cooperative efforts to achieve understanding at minimal conceptual and linguistic expense.

The Cooperative Principle implies that the act of verbalization as any other component of the speech act must be oriented towards the recipient in advance. In the process of language coding the speaker must provide the addressee with the effective cues for semantic text interpretation and thus facilitate the assimilation of its content by the listener.

It is supposed that any language has special means which function as orientation devices in the process of verbal decoding. These devices may not be relevant for expressing the informational content as such, but serve to organize it properly, giving it point and precision and thus contributing to its adequate understanding.

In comparison with other language functions the function of orientation has at least two peculiarities. The first one derives from the fact that the linguistic form of a speech utterance depends on the general communicative intention of the speaker and the speech situation.

As early as in I920-30-s some Soviet scholars pointed out, in particular, that monologue as a form of speech communication presupposing more or less prolonged speech impact on the listener typically conveys complicated conceptual information. As a result, monologue is bound to set greater demands to the act of verbalization than dialogue. In connection with this a well-known Soviet scientist L.P.Jakubinsky wrote:
"Монолог не только подразумевает адекватность выражающих средств данному психологическому состоянию, но выдвигает как нечто самостоятельное именно расположение, компонирование речевых единиц ... . Здесь сами речевые отношения становятся ... источниками появляющихся в сознании по поводу них переживаний... . На этой же почве возникают всевозможные явления синтаксического параллелизма и симметрии, т.к. сложность естественно вызывает какую-то организацию, построение" /2, p. 37/.

The second peculiarity consists in the fact that the use of language orientation means depends on the speaker's cooperative effort and on his evaluation of the intelligibility of his own utterance. As a consequence, orientation means are not subject to grammaticalization to the same extent as the means expressing the propositi-

Sy 1.5.1

onal context of a speech message.

To sum up, orientation means are related to such characteristics of the speech act as the complexity of information transmitted, time of speech impact, the listener's cognitive and cultural background and the current state of his mind, the degree of the speaker's cooperative effort and so on. All of these features point to a clearly pragmatic character of language orientation means.

Let's now turn to intonational phrasing. It is commonly agreed that the role of IP is more significant in monologue than in conversation form of speech. It is most actively used in texts that are complicated from the intellectual point of view and well-formed linguistically. In the following parts of this paper we are going to consider the ways in which IP can be used as an orientation device in the process of speech decoding. We'll also discuss the results of a psychological experiment that was undertaken with the purpose of illuminating of IP and revealing its functions in speech communication.

INTONATIONAL PHRASING AS A MEANS OF SENSE PROCESSING CONTROL IN SPEECH DECODING

It is a known fact that in a well-formed text IP correlates with its syntactic /in a wide sense/ structure. This is why IP presents a special interest for modeling speech decoding process. At the same time it's evident that IP is not the only possible means of transmitting information about the syntactic /deep and surface/ structure of a speech message. Syntactic relations between words, their correspondence to the general communicative intention of a speaker can be identified through morpho-syntactic cues /word order, form-words and morphemes/, lexical cues, verbal and situational context, the person's cognitive background and so on.

On account of such plurality of language means it seems justified to evaluate the contribution of each separate cue in the general procedure of sense decoding depending on whether understanding is possible in the absence of this particular cue. It is evident that IP /at least in Russian/ doesn't occupy a predominant position in the set of linguistic cues for decoding of speech message. It can be proved by various facts among which the following should be mentioned as the most significant:

1/ in speech reality utterances whose ambiguity is solved by IP solely are quite rare /as in "КАЗНИТЬ НЕЛЬЗЯ ПОМИЛОВАТЬ"/;

2/ written texts with no formal segmentation marks can be perfectly understood: for example, Old Church-Slavonic and Old Russian manuscripts which are known to have neither punctuation marks, nor word gaps or capitals;

3/ unnatural word-by-word text pronunciation makes its understanding more difficult but not impossible.

From what has been stated above it follows that IP reduplicates the function of the lexical-syntactical cues that are constantly present in any written and oral utterance. IP can be regarded then as one of non-obligatory, redundant

linguistic devices serving to increase the reliability of sense decoding in oral communication. This conclusion, however, doesn't take account of the most important characteristics of IP, and namely, the fact that IP separates speech currently and at the same time takes part in forming it as rhythmically organized time process.

Following many Soviet linguists and scholars in literature we consider the speech to be rhythmically organized if it is divided into the subjectively isochronic speech stretches with utterly or partly reproducible phonetic structure. IP's correlation with speech rhythm is recognized by the majority of Soviet phoneticians and there is no necessity to prove it here. At the same time we'd like to discuss some peculiarities of IP in its relation to rhythm from the point of view of speech analysis-synthesis process.

Considering speech generating process it is usually noted that rhythm as one of IP's forming factors is functioning under constant sense control /or corrected intellectually/. This idea is often accompanied with that of considering IP as some means of packaging the lexical content of an utterance into some rhythmically organized linguistic form. This is hardly true. "A cognitive draft" /a well turned expression of L.P.Jakubinsky/ becomes the verbal text on account of the simultaneous acting of semantic, pragmatic and phonorhythmic factors. Thanks to this the generated text meets both the needs of adequate transmission of information and needs of comfortable pronouncing and perception.

From the point of view of speech perception some ideas of contemporary cognitive psychology are very interesting. In particular it advances an assumption that solution of different intellectual tasks, speech recognition among them, is performed by a listener on the continuously acting rhythmical background which regulates his attention, memory and other cognitive devices. It is possible that in the case of speech perception rather quick and adequate detection of IP's units is provided not only by listener's knowledge of the general rhythmical principle of speech patterning but also due to his own rhythmical activity getting him into the state of pre-expectation or preparedness for detection of certain fragments in rhythmically organized text he hears. In this case the effort which the listener expends on sense decoding depends on the extent to which the speaker coordinates information structure of the message with its intonational phrasing. Distinctness of IP's phonetical realization and correlation of the speaker-listener rhythmical activity are also important.

The rhythm-forming character of IP permits us to speak of its role in the process of speech decoding in the following way. On the one hand, in view of IP's correlation with syntactic structure of an utterance we can consider the current detection of rhythmo-intonational units to be the first step of sense decoding proper. On the other hand, IP is set by the speaker and we can see that by generating it the speaker is governing the procedure of listener's speech perception. The prosodic means with the help of

which IP is realized can be regarded then as special devices whose function is to orientate the listener in sense space of the text developing in time. Everything that has been told above let us consider IP to be a direct result of the cooperative principle in speech communication.

EXPERIMENT

Design

The aforesaid considerations must be undoubtedly supported by some experimental ground. To our regret there are practically no concrete data about the ways of sense decoding and we do not know how IP is used in this process. It seems reasonable that in such situation first of all it is necessary to reveal the most striking features of sense processing of a text which can have two different forms of presentation:1/ with no formal segmentation marks, 2/ naturally pronounced. We have conducted the pilot study in this direction and we'll outline the results below.

For the experiment we have chosen a sample of a scientific text in Russian. The length of the text was about 300 words.

Procedure

The study was divided into two stages. On the first stage the subjects /12 persons/ were given the text in the written form, printed in one line without word gaps, punctuation marks and capitals. We'll call it further "written text" /WT/. On the second stage /about a month later/ the same subjects were presented the same text but naturally pronounced. We'll call it further "naturally taken text" /NTT/.

On both stages the subjects must have carried out the same task: to rewrite the analysing text without word alteration in the most convenient and usual for them form. The subjects were informed beforehand that the text was rather lengthy and they have to copy it piece-by-piece. Some restrictions on the procedure of selecting the fragments were fixed:

1/ In selecting a fragment next in turn the subjects were allowed to read or to listen the text only once starting from the point which was defined by the endpoint of the preceding fragment.

2/ The subjects were not permitted to read the text aloud /or to repronounce it/. It is of interest to note that, as the subjects admitted themselves later, they nevertheless articulated the WT silently.

3/ The choice of a rewritten fragment was claimed to be done under sense control, that is a chosen fragment is supposed to stand out as some sense unit including a word or a succession of words. Besides these there were no strict restrictions regarding the fragment length and the concrete character of sense unity. In accordance with the general principle of sense control the subjects were free to choose any fragment best suitable for them to carry out the task set.

The goal of the experiment reported here was to compare the subjects' behaviour in analy-

sing the WT and that of NTT. For this purpose some formal characteristics have been fixed for every fragment chosen by each subject: a) time of selection /in sec/, b) length /in words/, c) the presence of the so-called "looking ahead" or a situation when the chosen fragment was smaller than the text piece that had been read or listened for its selection, d) necessity of correcting the place of the preceding fragment's endpoint or a situation when a subject selecting a fragment coming in turn discovers that he was wrong in the choice of the endpoint of the preceding fragment, e) the presence of word-alterations in the fragment written down /errors in word identification, omissions, substitutions/.

In addition to the above mentioned characteristics we have also fixed some syntactic properties of the fragments written down by every subject. These properties can be called indicators of the fragments' coherence, completeness and autonomy.

The features in question can be analysed and formally determined only in case of fixed representation of syntactic structure of an utterance. The method of representation applied in our study is known in syntax /3/ as the linearized syntactic dependency tree /LSDT/. This kind of representation takes account of both syntactic relations between words and word order.

Briefly, these features are as follows:
- coherence. A fragment is considered to be coherent if all the words within it are linked, directly or indirectly, and only for one word the head is out of the given fragment. Otherwise the analysing fragment is considered to be incoherent. Formally coherence means that the only one external arrow of LSDT can enter the fragment.
- Completeness. A fragment is considered to be complete if all the words within it have no subordinates out of the fragment. A fragment is considered to be partly incomplete if the only word in it which has subordinates out of the given fragment is the main predicate of an utterance. And finally, a fragment is considered to be incomplete if there is at least one word being not the main predicate of the utterance has a subordinate out of the given fragment. Formally completeness means that there is no external arrow coming out of the fragment.
- Autonomy. A fragment is considered to be autonomic if it corresponds to a simple sentence or, in other words, if it is equal to the entire argument-predicate structure. Otherwise the analysing fragment is considered to be non-autonomic. Formally autonomy means that the only external arrow which enters the given fragment is that to the symbol S of LSDT. The latter occurs only in cases of syntactically compound utterances.

For the experimental goal it was interesting to study the behaviour of different subjects writing down the text presented in the same form /WT or NTT/ and behaviour of the same subject writing down the text presented in different forms /WT and NTT/. To estimate the homogeneous character of subjects' behaviour we have calcu-

lated the relative number of coinciding fragments for each pair of experimental text copies. It seemed also useful to value the stability of segmentation locations /i.e. boundaries of selected fragments/. Numerically it can be expressed as the relative amount of the subjects which have distinguished the given boundary. The frequency distributions of boundaries with different stability can serve as a measure of behavioral homogeneity of the whole group of subjects having analysed the same text.

Results and analysis

The conducted experiment has exposed many interesting peculiarities in the subjects' behaviour. Here we have to discuss them in a brief and conspective form.

First of all it is necessary to note that the rewriting of the WT /text with no formal segmentation marks/ was admitted by all the subjects as a rather complicated task which required a great intellectual and psychological effort. According to the strategy of text rewriting all the subjects can be subdivided into two groups. The main distinctive feature of this division is the preferable length and syntactic properties of the fragments having been singled out. It seems that the basis of this distinction is the difference in the volume of subjects' short-term memory. It should be noted that all the subjects were approximately of the same age /25-35 years old/ and the same social rank /collaborators of the Moscow University philological department/. From the whole group /I2 persons/ the majority of subjects /7/ preferred to select rather short fragments, 3 persons, on the contrary, singled out rather long stretches, and the behaviour of 2 persons was inconsistent. The results of the latter group were not considered later.

We'll denote further the first group of subjects as I-G and the second one as II-G.

The ordered set of fragments, i.e. the copied text written down by every subject, was analysed from the point of view of the formal characteristics described above. Table I presents the results of this analysis.

Table I shows that the change of the text presentation form from WT to NTT provokes the following modifications of SF characteristics:
I. The length of SF becomes at an average one and a half longer, what causes decreasing the total number of SF in subjects' NTT copies on 31,4%.
2. The time expended on fragment selection shortens twice.
3. The number of fragments singled out with the "looking ahead" strategy substantially diminishes; it is seen that for the subjects having used this strategy in writing down the WT the number of SF /with "looking ahead"/ makes 45% of the total. In writing down the NTT the same subjects used this strategy only in 2-5% of all the cases.
4. The need of current correction of SF boundaries drops utterly.
5. The number of SF with word alteration is rather small in both cases /6-7% for I-G and I6-

Table I. Characteristics of selected fragments /SF/. Mean values obtained for all the subjects of the same group. Values in % are given in relation to the total number of SF.

	WT		NTT	
	I-G	II-G	I-G	II-G
Length of SF	3,I	5,2	4,6	7,7
Time of selection	9,2	I2,5	4,3	6,4
The total SF number	86	5I	59	35
Number of F selected with "looking ahead"	8,9%	44,8%	6,5%	2,I%
Number of SF with correction	9,0%	6,I%	-	-
Number of SF with word alteration	6,4%	I6,0%	7,I%	22,7%
Number of incoherent SF	8,2%	9,7%	8,5%	2,6%
Number of complete SF	44,5%	5I,6%	49,4%	56,7%
Number of partly incomplete SF	I4,2%	8,8%	I2,9%	I,9%
Number of incomplete SF	41,3%	39,3%	37,6%	40,8%
Number of autonomic SF	7,0%	30,4%	22,3%	46,9%

23% for II-G, the greater value for the latter clearly correlates with more lengthy SF/.
6. As for the syntactic indicators the autonomy modificates most of all: in NTT copies the number of autonomic SF increases essentially /at an average I5-I7%, for some subjects up to 50%/. The coherence and completeness indicators change insignificantly.

The character of the syntactic indicators' modifications and also that of the SF length allows us to conclude that the diversity in the same subject's behaviour as to WT and NTT cases is not based on differences in text comprehension by the speaker and by the subject himself. Most likely, the truth is that the speaker in his producing the text sets the higher level of syntactic structure of an utterance, i.e. the level being nearer to the entire sense frame of a message.

Let's now discuss the homogeneity of subjects' behaviour. As an illustration, the data concerning the relative number of coinciding

---

fragments in text copies for some subjects are presented in table 2.

Table 2. The number of coinciding fragments in text copies of different subjects /in % of the total number of F in each pair of copies/

SUBJ.		WT			
		I	I-G 2	3	II-G I
I-G	I	38,6	56,8	54,7	26,3
N T T	2	63,9	28,0	55,7	24,6
	3	70,I	64,8	29,0	26,I
II-G	I	68,3	6I,3	68,5	67,9

In table 2 the diagonal corresponds to the same subject in case of WT and NTT. Above the diagonal there are results for the different subjects in WT case, under it - the results in NTT case.

Table 2 shows that in NTT case the homogeneity of selecting the fragments increases significantly. It is expressed by the difference obliteration between two groups of subjects and especially between various subjects in the same group. Thus, for example, the analysis reveals that in I-G the relative number of coinciding fragments is about 55% for a pair of subjects in WT case, but about 72% in NTT case. At the same time in different text copies /WT and NTT/ written down by the same subject there are only 37,5% of coinciding fragments.

As to the stability of segmentation locations the data concerning this problem are represented on fig. I.
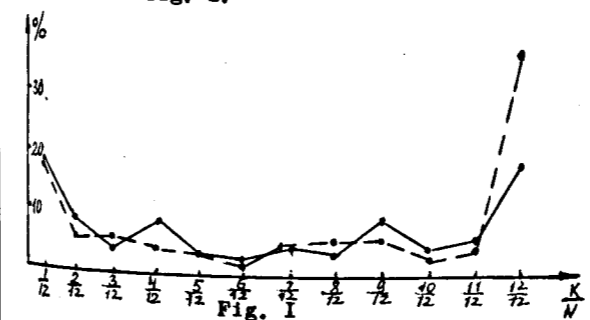

Frequency /in % / distribution of boundaries with different stability. Continuous and dotted lines correspond to WT and NTT cases accordingly. $\frac{K}{N}$ - the relative number of subjects having distinguished the given boundary.

Our study reveals that in NTT case the total number of boundaries distinguished by the whole group of subjects decreases greatly in comparison with WT case /from I34 to 85/. At the same time there is the increasing of the amount of boundaries distinguished by all the subjects, i.e. the most stable ones /from I8% to 35%/.

CONCLUSION

In this study we proceeded from the following assumptions:
I. The process of piece-by-piece rewriting of a text realizing under sense control reflects to some extent the real process of sense decoding.
2. A "written" text with no segmentation marks is equivalent to its pronounced copy from the point of view of sound information included and the only difference between the copies is the absence of prosodic cues in WT.

If these assumptions are reasonable to some degree, then the obtained results permit us to make the following conclusion: IP, being the main determining factor of prosodic characteristics of the analysed text, controls the strategy of sense processing, unifies it and makes it more optimal from its aim point - decoding of the entire sense frame of a message. It is scarcely probable that the diversity of subjects' behaviour in WT and NTT cases is explained only by different perception modes and in no connection with IP cues.

It stands to reason that further research is necessary. In particular, it is interesting to compare our results with the analysis of subjects' behaviour in case of both a normally organized written text and a text pronounced monotonically word-by-word.

REFERENCES

/I/ Grice H.P. Logic and conversation. - In: Syntax and Semantics, v.3. N.Y. Ac.Press, I975, p. 4I-88.

/2/ Якубинский Л.П. Избранные работы. Язык и его функционирование. М., I986.

/3/ Падучева Е.В. О способах представления синтаксической структуры предложения. - ВЯ, I964, № 2.