

On the acquisition of the probabilistic prediction strategy

MARINA RUSAKOVA

State Pedagogical University of Russia, St. Petersburg

INTRODUCTION

Normal speech processing is provided to a great extent by the strategy of probabilistic prediction. The usage of this strategy would be impossible if the objective frequencies of different speech units were not represented in psycholinguistic organization.

It is known that the system of subjective probabilities reflects the real statistic organization of speech at the same time differing from it in some essential features. A lot of experiments shows that metalinguistic competency of native speakers includes the ability to evaluate objective frequencies of words and this evaluation corresponds to word perceptive and articulative characteristics (Frumkina, 1971). Language units with the high subjective probability are produced and comprehended not absolutely the same way as that of with the low subjective probability.

Perceptive and articulative characteristics of different language units depend, thus, more on their subjective frequencies than on the objective distribution of this units in speech.

The *item* of the present paper is to find whether the strategy of probabilistic prediction is used in grammatical component of speech production and to determine some regularities in the acquisition of the ability to use this strategy by Russian children. The system of Russian nominal inflection was investigated from this point of view.

There are 6 cases in Russian and two numbers. Each flexion is a marker of case and number, thus, we have 12 case-number forms.

The very fact that the system of subjective frequencies exists as a part of language competence and used in speech processes is under the question.

EXPERIMENTS, METHODS, SUBJECTS AND DATA

The first part of the paper concerns the distribution of the case-number forms in child and adult speech.

The distribution of the case-number forms in the speech of a 2.5-year old Russian child (the corpus of texts consisted of nearly 40 000 words) occurred to be nearly the same as that of the adult speech, while there were yet a lot of different grammatical mistakes in using such forms by the child. It means that the child does not avoid the difficult forms and the distribution depends only on the relations of the objects of the external world, this relations are thus acquired very early. So, the base to acquire the predicting strategy also appears very early.

The other part of the present paper concerns the comparison of the subjective frequencies system of the elder children (60 11-year olds took part in the experiment) and that of the adults (200 adults took part in the experiment). Subjective frequencies were determined by the

method of scaling. The informants (both adults and children) had to evaluate the frequency of every case-number form, giving them marks from one (which means a very rare form) to seven (which means a very frequent form). The results of this experiment are shown in the tables. First of all it should be mentioned that only few informants refused to fulfill the task, the others didn't have any difficulties, thus showing that frequencies of the 12 case-number forms are this or that way represented in their linguistic competence. Table 1 and table 2 show the results of this scaling. Table 5 shows also that adults were rather unanimous in the evaluation of frequencies of some forms, especially in the evaluating of the nominative case frequency: 61% of the informants gave this form the mark, corresponding to the median of the distribution. The unanimity of the answers and the high positive correlation between the results of scaling and perceptive characteristics (received in a special experiment which we can not describe just now) of case-number forms show that the probabilistic predicting strategy is used by adults in case-number forms processing.

Table 3 and table 4 show the correspondence of the subjective frequency system to the objective one. It is seen in the table two that the correlation of subjective and objective orderly patterns is very high, but it is not the full coincidence. It should be also added and this is not seen in the table that informants think the case forms to be less different in their frequencies than they are in reality. The main feature of the subjective system is the contrast of the nominative and the oblique cases. The contrasts within the oblique cases are not felt strictly by the language speakers. So, the subjective system of case forms frequencies is more simple and more generalized than the real one. Table 4 shows that if we concern not only case, but case-number forms the correlation of the subjective and the objective systems will not be so high. The informants think the number forms to be much more alike than they are in reality. In fact we have two independent case frequency systems: one for the singulars and one for the plurals. This fact can be probably explained if we admit that processing of the number forms in Russian is more the word formation than the inflection, then we don't use probabilistic prediction strategy in this grammar component and don't need the subjective frequency system.

The last lines of all the tables show that subjective frequencies system of the 11-year olds differs from that of the adults:

1. It is less simple in its structure;
2. It looks more like the real statistical distribution of such forms in speech;
3. It is more oriented to the colloquial speech. In Russian different styles of speech are characterized by different distribution of genitive and accusative. In all literal forms of speech genitive is more frequent than accusative;
4. Table 5 shows that children have two picks of unanimity: nominative and accusative, it means that the main generalization of adults nominative oblique is not yet acquired by children.

DISCUSSION

The main conclusions of the present investigation are:

1. Native language speakers use the probabilistic prediction strategy in speech processing not only within word component, but in the grammar component also. Frequency system investigation can help to describe some internal psycholinguistic mechanisms.

2. Statistic organization of nominal inflection is formed very early (while statistic organization of other speech production components is not), thus, the possibility to form the system of subjective frequencies supplying the strategy of the probabilistic predicting appears very early too. Nevertheless, even with the elder children the subjective frequencies system is not completely formed. Thus, the capacity to use the probabilistic predicting strategy develops for a very long time and represents one of the subtle distinctions in speech capacity of elder children from that of adults.

REFERENCES

Frumkina R. M. ed. 1971. *Verojatnostnoje prognozirovanije v rechi*. Moscow, Nauka.

TABLES

Table 1. MEDIANS OF EVALUATION DATA SET

		Nom.	Gen.	Dat.	Acc.	Instr.	Prep.
adults	sing.	7	6	5	5	3	4
	plur.	7	5	4	4	3	4
children	sing.	7	6	5	6	5	5
	plur.	6	5	4	5	3	3

Table 2. ORDERLY PATTERN OF CASE FORMS SUBJECTIVE FREQUENCIES (DESCENDING)

adults	sing.	Nom.	Gen.	Acc.	Dat.	Prep.	Intr.
	plur.	Nom.	Gen.	Acc.	Prep.	Dat.	Instr.
children	sing.	Nom.	Acc.	Gen.	Dat.	Instr.	Prep.
	plur.	Nom.	Acc.	Gen.	Dat.	Instr.	Prep.
rank		1	2	3	4	5	6

Table 3. RANKS OF SUBJECTIVE AND OBJECTIVE FREQUENCIES OF CASE FORMS

objective		Nom.	Acc.	Gen.	Prep.	Instr.	Dat.
subjective	adults	Nom.	Gen.	Acc.	Prep.	Dat.	Instr.
	children	Nom.	Acc.	Gen.	Dat.	Prep.	Instr.

Table 4. ORDERLY PATTERN OF CASE AND NUMBER FORMS FREQUENCIES (DESCENDING)

Obj.		Nom. Sing.	Acc. Sing.	Gen. Sing.	Prep. Sing.	Nom. Plur.	Instr. Sing. & Acc.	Gen. Plur.	Dat. Sing.	Prep. Plur.	Instr. Plur.	Dat. plur.
Subj.	Adult	Nom. Sing.	Nom. Plur.	Gen. Sing.	Acc. Sing.	Dat. Sing. & Gen. Plur.		Prep. Sing. & Acc. Plur. & Prep. Plur.			Dat. Plur.	Instr. Sing. & Instr. Plur.
	Child	Nom. Sing.										

Table 5. UNANIMITY IN EVALUATION OF FREQUENCY (%)

Adults						
<i>Singular</i>						
Nom.	Gen.	Dat.	Acc.	Instr.	Prep.	
64	30	24	19	30	21	
<i>Plural</i>						
Nom.	Gen.	Dat.	Acc.	Instr.	Prep.	
51	20	19	19	19	17	
Children						
<i>Singular</i>						
Nom.	Gen.	Dat.	Acc.	Instr.	Prep.	
61	28	28	19	23	18	
<i>Plural</i>						
Nom.	Gen.	Dat.	Acc.	Instr.	Prep.	
29	19	24	36	29	24	

SUBJECTIVE FREQUENCIES OF NUMBER FORMS OF FREQUENT AND RARE NOUNS

Table 6. Adults		
	Frequent	Rare
Sing.	242	257
Plur.	221	153

Table 7. Children		
	Frequent	Rare
Sing.	1156	1034
Plur.	1134	1052