The primary objective of this paper is to describe an experiment designed to investigate the semantic relationships between the three basis components of a prepositional construct: the governor, preposition and the complement. Because of the preliminary nature of the experiment, only simple data processing equipment, such as the keypunch and the sorter, was used. The implementation of this approach on a larger scale, however, would necessitate the use of more sophisticated hardware.

The described procedure uses Russian prepositions because, while working on this problem, the author was a research staff member of the Russian-English mechanical translation group at IBM's Thomas J. Watson Research Center in Yorktown Heights, New York.

While the described procedure presents a tentative approach, which does not offer a solution to the semantic ambiguities within prepositional constructs in Russian, it does suggest a method for examining each basic component of a given construct in relation to other constructs containing different types of prepositions.

The data used in the model was collected mainly from the Soviet Academy of Sciences Grammar and, to some extent, from the Soviet Academy of Sciences Dictionary. Initially an attempt was also made to compile data from other dictionaries. It was found, however, that the presentation and the classification of the data was not detailed enough for the purposes of this study. Therefore, only some of the

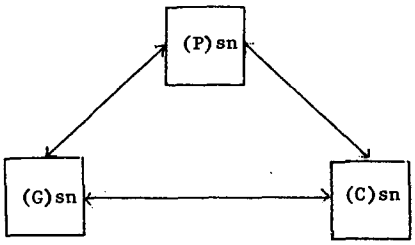
prepositions not listed as such in the previously named sources were included in the experiment. The next logical step, using the arrangement of the data as shown below, should be the culling out of additional data in the case of Russian, and complete data in the case of other languages, from dictionaries, concordances and random texts. Following various sorting patterns, the results should then be tested through generative processes and checked against concorded 'real life' examples.

General Purpose

As stated earlier, the purpose of the proposed approach is the establishment of patterns of semantic correlations between:

1. Given Governor and its Preposition G⟷P (left boundaries)
2. Given Preposition and its Complement P⟷C (right boundaries)
3. Given Governor and its Preposition's Complement G⟷C

These relationships can be diagrammed as follows:



where sn=sematic property of any value.

If either of the semantic components is found to exclusively govern the combination of the two remaining semantic components then it can be said that

(G)sn + (C)sn + (P)sn = Sx

where

Sx = valid semantic pattern = sum of semantic properties of (G)sn(C)sn and (P)sn

Where confirmed, the established semantic patterns can be incorporated into a look-up table as an additional tool for parsing procedures and for testing the validity of class-membership within prepositional phrases. Where not confirmed or where the patterns are not clear, further refinement of the semantic properties of the components in question must be carried out.

Method of Procedure

Following the classification provided by the Academy of Sciences Grammar (henceforth referred to only as Grammar), prepositional constructs were coded according to the grammatical category of their governors and the type of relationship (henceforth referred to only as TR) with their complement. The following governor categories were recognized: noun, verb, adjective, adverb, numeral.

Next, the governors as well as the complements were coded according to their semantic properties from examples provided by the

Grammar. This semantic classification was, in a sense, arbitrary in that it did not attempt to map the universe of all Russian words in the five governor categories, but tended to accomodate only those complements and governors contained in the examples, and only if their characteristics seemed relevant. Thus, if the Grammar provided two examples of the usage of the preposition and the semantic properties of respective governors and complements were clearly identical, only one example was selected. The listing of these constructs in the five governor categories is given in Appendix I.

Since, in an initial study of this type, a large number of semantic classes might tend to obscure the existence of possible patterns, an attempt was made to keep the number of these classes at a minimum. As stated earlier, the adoption of this approach in an extensive study of constituents within prepositional constructs would require more elaborate semantic mapping. For the purposes of this study, the total number of semantic classes for nouns was narrowed to 24, for verbs 9 and 6 for adjectives. (See Appendix II)

The classification of numerals and adverbs as governors was abandoned when it was discovered that, according to the examples provided by the Grammar, their semantic values in no instance effected the selection of a complement of a given semantic category. Thus the

examples citing the usage of cardinal, ordinal and indeterminate numbers such as "НЕСКОЛЬКО" showed that these governors may co-occur with a complement of any class. A similiar phenomenon was observed in the behavior of adverbs acting as governors. It is possible that a more detailed study of a large corpus will reveal the existence of definite relationships.

The total number of types of relationships (TR) abstracted from the Grammar was 42 (see Appendix III). Since some of the prepositions, especially of the compound or adverbial type, were not provided by the Grammar with an example of their usage in a given TR, the latter was derived through transformational cross-reference from the Academy of Sciences Dictionary, and when that proved impossible, it was supplied by the native speaker of Russian employed by the Project.

In order to fit the data for each TR on a single IBM card (for easier sorting), those TRs which seemed somewhat redundant or insufficiently documented were combined and the total number of TRs was reduced to 11. Again, while the TRs were translated literally from the Grammar (admitedly, some of the translations seem a little awkward, e. g. 'togetherness'), the reduction of their total number was an arbitrary arrangement aimed at simplifying the overall research procedure. The manner in which the 43 TRs were reduced to 11 is shown in Appendix IV.

The 11 TRs used in the final analysis are as follows:

1. Attributive	7. TR expressing togetherness
2. Temporal	8. similarity
3. Spatial	9. deprivation
4. Purpose	10. conformity
5. Causative	11. comparison
6. Objective	

In addition to the aforementioned, the following data was added to every construct:

1. General relationship

AA purely attributive

OO * objective

MM * modifying

XX * other

AO attributive with objective shading

OA objective with attributive shading

etc.

2. Relative position to the complement

A PNP precedes the governor

P PNP follows the governor

X no preference

3. Structural restrictions

X none

I preposition-complement form an idiom in a given configuration

preposition (e.g. СОКРАТИЛОСЬ В НЕСКОЛЬКО РАЗ)

P governor - preposition form an idiom in a given configuration

e.g. (перевёл по рангу)

4. Equivalence with other TRs (syntactic difference - semantic equivalence)

1 yes

0 none

5. Other transformational possibilities

1 yes (e.g. сел за стол - сел у стола)

0 none

6. Letter codes to facilitate file search

(usually the first three letters of the preposition)

When punched on IBM card the data had the following format:

Columns	Data
3-23	preposition
25	case governed by the preposition in a given TR
27-31	specific relationship with the complement
33-34	general relationship with the complement
36	relative position to the complement
38	governor class
40-41	governor sub-class
43	structural restrictions
45-46	complement class

(cont.)

Columns	Data
48	equivalence with other constructs
50	other transformational possibilities

The final step consisted of sorting the cards on individual columns for detection of patterns.

APPENDIX I

Preposition	Noun as governor	Verb as governor	Adjective as governor	Adverb as governor	Numeral as governor
Bez	(NA↓A)=AT	(VX↓K)=AT (VX↓B)=OB			
Blagodar4*		(VX↓A)=CA			
Bliz	(NF↓B)=AT				
Cerez	(NM↓Q)=AT	(VX↓B)=OB (VX↓U)=SP (V4↓E)=SP (VX↓R)=TE			
D14	(NE↓E)=AT (NK↓B)=AT	(VX↓C)=OB (VX↓A)=PU	(A2↓A)=CO (A2↓L)=OB		
Do	(NC↓K)=AT (NC↓E)=AT (NP↓E)=AT (NE↓B)=AT (NL↓B)=SP (NL↓Q)=TE	(VX↓Y)=AT (VX↓A)=OB (VX↓B)=SP (VX↓Q)=TE	(AX↓Y)=AT (A3↓Y)=OB (A1↓E)=SP (A5↓Q)=TE	(DX↓S)=TE	
Iz	(NC↓A)=AT (NF↓B)=AT (NA↓E)=SP	(VX↓W)=AT (VX↓V)=CA (VX↓E)=OB	(AX↓B)=AT	(DX↓A)=SP	(RX↓B)=AT

To be read as follows:

(grammatical and semantic categories of governor ↓ semantic category of complement) = Type of Relationship

*4 = Я

Appendix I (cont.)

(cont.)

		(VX→G)=OB		
		(VX→B)=OB		
		(V1→H)=SP		
		(VX→K)=OB		
		(V3→B)=SP		
Iz-pod	(NF→D)=AT	(VX→A)=SP		
	(NE→F)=AT			
	(NL→E)=AT			
Iz-za		(VX→B)=CA		
		(VX→K)=CA		
K	(NI→K)=AT	(VX→B)=OB	(AX→K)=AT	(DX→B)=SP
	(NI→B)=AT	(VX→A)=OB	(AX→A)=OB	
	(NK→E)=AT	(V7→K)=PU	(AX→Y)=SP	
	(NE→B)=AT	(VX→Q)=TE		
	(NL→A)=OB			
	(NP→A)=OB			
Krugom	(NF→B)=AT	(VX→B)=SP	(AX→B)=SP	
Mejdu*	(NA→A)=AT	(VX→Q)=TE	(AX→D)=TO	
	(NL→B)=SP	(VX→D)=TO		
	(NK→A)=TO			
Mimo		(VX→B)=SP		
Na acc.	(NB→A)=AT	(VX→Y)=AT	(AX→A)=AT	
	(NL→A)=AT	(VX→F)=AT	(AX→Y)=AT	
	(NI→H)=AT	(VX→V)=AT	(AX→B)=OB	
	(NC→E)=OB	(VX→E)=AT		
	(NL→A)=OB	(VX→A)=OB		
	(NL→K)=PU	(V8→A)=PU		
	(NL→B)=SP	(VX→B)=SP		
	(NN→A)=SP	(VX→H)=SP		
	(NM→H)=SP	(VX→R)=TE		
	(NO→B)=SP			

*j = ѓ

Appendix I (cont.)

Na prep.	(NB→E)=AT (NL→B)=AT (NL→Q)=AT (NL→E)=OB	(V2→E)=AT (VX→Y)=AT (V2→F)=OB (VX→A)=OB (VX→K)=SP (VX→F)=SP (VX→J)=SP (VX→E)=SP (VX→Q)=TE	(AX→Y)=AT (AX→E)=SP (AX→S)=TE		(RX→E)=AT
Nad	(NL→B)=AT (NL→Y)=OB	(VX→Y)=OB (VX→B)=SP		(DX→B)=SP	
Naperekor		(VX→A)=OB			
Naprotiv		(VX→B)=SP			
Nascet*	(NE→A)=AT				
Navstrecu		(VX→L)=SP (VX→B)=SP			
Vnutri		(VX→B)=SP			
O acc.	(NL→B)=OB	(VX→E)=OB			
O prep.	(NY→A)=AT	(VX→A)=OB			
Okolo	(NF→B)=AT	(VX→B)=SP	(AX→E)=SP		
OT	(NE→B)=AT (NB→B)=AT (NK→Q)=AT (NE→A)=AT (NL→E)=AT (NE→E)=AT	(VX→K)=CA (VX→A)=OB (VX→H)=SP	(AX→Y)=CA (AX→A)=DE (AX→B)=SP (AX→Q)=TE	(DX→B)=SP	(RX→B)=AT
Otnositel6no	(NE→A)=AT				

*c = 4

Appendix I (cont.)

Pered	(NB+B)=AT	(VX+A)=OB	(AX+A)=AT	(DX+M)=TE	
	(NL+B)=AT	(V2+A)=SP	(AX+B)=OB		
	(NL+Q)=AT	(VX+Q)=TE	(AX+S)=TE		
	(NR+S)=AT				
	(NK+B)=OB				
	(NL+A)=OB				
Po acc.	(NB+B)=AT	(VX+E)=SP	(AX+E)=SP		
	(NA+A)=AT	(VX+Y)=AT	(AX+E)=AT		
	(NJ+Y)=AT	(VX+K)=CA	(AX+Q)=TE		
		(VX+B)=OB			
		(VX+H)=SP			
		(VX+A)=SP			
		(VX+Q)=TE			
Po prep.		(VX+A)=OB		(DX+B)=SP	(RX+E)=AT
		(VX+L)=TE			
Po casti		(VX+K)=OB			
Po napravlenih k*		(VX+B)-SP			
Po otnowenih k		(VX+B)=OB			
Po povodu		(VX+A)=CA			
Po pricine		(VX+Y)=CA	(AX+Y)=CA		
Po slucah		(VX+A)=CA			
Po sravnenih s			(AX+A)=CP		
Pod acc.	(NB+E)=AT	(VX+L)=AT			
	(NE+Y)=AT	(VX+A)=OB			
		(VX+E)=SP			
		(V2+L)=OB			
		(VX+B)=SP			
		(VX+R)=TE			

*h = ĸ

Appendix I (cont.)

Pod instr.	(NF+F)=AT	(VX+Y)=AT	(AX+E)=SP	(DX+B)=SP
	(NL+K)=AT	(VX+K)=OB		
	(NE+E)=AT	(V6+K)=SP		
	(NL+E)=SP	(VX+J)=SP		
		(VX+E)=SP		
Podle	(NF+B)=AT	(VX+B)=SP	(AX+E)=SP	
Podobno		(VX+A)=SI		
Poperek		(VX+K)=OB		
Posle		(VX+A)=TE	(AX+Y)=TE	
Posredstvom		(VX+K)=AT		
Poverx*	(NF+B)=AT	(VX+B)=SP		
Pozadi		(VX+B)=SP		
Prejde		(VX+B)=TE		
Pri		(VX+K)=AT	(AX+A)=AT	
		(VX+E)=SP	(AX+K)=TE	
		(VX+D)=TE		
Pro		(VX+A)=OB		
		(VX+K)=PU		
Protiv	(NB+K)=AT	(VX+A)=OB		
	(NF-B)=AT			
	(NL+B)=AT			
Putem		(VX+K)=AT		
R4dom s			(AX+B)=AT	
Radi		(VX+C)=OB		
S acc		(VX+T)=AT		

*x = X

Appendix I (cont.)

S gen.	(NL→K)=AT	(VX→K)=CA	(AX→E)=SP	
	(NL→B)=AT	(VX→L)=OB	(AX→Y)=CA	
	(NB→E)=AT	(VX→B)=SP	(AX→Q)=TE	
		(VX→Q)=TE		
S instr.	(NC→K)=AT	(VX→B)=AT	(AX→E)=AT	(DX→B)=OB
	(NF→B)=AT	(VX→V)=AT	(AX→B)=AT	
	(NF→K)=AT	(VX→K)=OB	(A4→B)=SP	
	(NB→A)=AT	(V2→B)=OB		
	(NE→E)=AT	(VX→E)=OB		
	(NC→B)=AT	(VX→A)=OB		
	(NX→B)=OB	(VX→Q)=TE		
	(NL→A)=OB	(VX→D)=TO		
	(NL→B)=AT			
	(NB→B)=TO			
Skvoz6*		(VX→A)=OB	(AX→B)=SP	
Soglasno		(VX→A)=SI		
soobrazno		(VX→Y)=SI		
Sootvestvenno		(VX→Y)=SI		
Sorazmerno		(VX→Y)=SI		
Sredi			(AX→E)=SP	
Szadi		(VX→B)=SP		
U	(NC→C)=AT	(V8→B)=OB	(AX→N)=SP	(RX→E)=AT
	(NC→E)=AT	(VX→B)=SP		
	(NL→B)=AT			
V acc.	(NF→E)=AT	(VX→K)=AT	(AX→G)=AT	
	(NF→T)=AT	(VX→A)=OB	(AX→L)=TE	
	(NL→E)=AT	(V1→K)=OB		
	(NY→T)=AT	(VX→B)=SP		
	(NF→B)=AT	(VX→Q)=TE		
	(NL→E)=OB			

*6 = б

Appendix I (cont.)

(cont.)	(NL→A)=PU (NL→B)=SP (NL→Q)=TE				
V prep.	(NB→E)=AT (NL→B)=AT (NA→B)=AT (NB→B)=OB (NL→B)=OB (NK→K)=OB	(VX→K)=AT (VX→E)=AT (V7→K)=OB (V8→K)=OB (VX→A)=OB (VX→E)=SP (VX→U)=SP (VX→Q)=TE	(AO→M)=AT (A3→N)=AT (AX→Y)=AT (AX→E)=SP	(DX→B)=SP	(RX→E)=AT
V dele		(VX→K)=OB			
V oblasti		(VX→K)=OB			
V otnowenih k			(AX→A)=AT		
V otnowenii		(VX→K)=OB			
V prodoljenie		(VX→Q)=TE			
V qel4x*		(VX→K)=PU			
V silu		(VX→Y)=CA			
V tecenie		(VX→Q)=TE			
Vblizi	(NF→B)=AT	(VX→B)=SP			
Vdol6	(NF→B)=AT	(VX→B)=SP			
Vmesto	(NA→A)=OB				
Vnutr6		(VX→B)=SP			
Vokrug	(NF→B)=AT	(VX→B)=SP			
Vopreki		(VX→A)=OB			

*q = ц

Appendix I (cont.)

Vozle	(NF→B)=AT	(VX→B)=SP	(AX→E)=SP	
Vperedi		(VX→B)=SP		
Vsled		(VX→B)=SP		
Vsledstvie		(VX→Y)=CA		
Vvidu		(VX→Y)=CA		
Vzamen	(NA→A)=OB			
Za acc.	(NL→K)=CA	(VX→A)=OB	(AX→A)=AT	
	(NL→A)=OB	(VX→B)=OB	(AX→Q)=TE	
	(NL→R)=AT	(VX→Y)=OB		
	(NL→J)=SP	(VX→B)=SP		
	(NL→R)=TE	(VX→Q)=TE		
Za instr.	(NB→B)=AT	(VX→B)=AT	(AX→B)=SP	(RX→E)=AT
	(NL→K)=AT	(V2→B)=OB		
	(NL→E)=AT	(V5→B)=SP		
	(NL→A)=OB	(VX→A)=OB		
		(VX→S)=TE		

APPENDIX II

Nouns

Any noun	A
Concrete	B
animate	C
personal	D
inanimate	E
objects	F
mass	G
location	H
linear (river, road, valley, etc.)	I
geographical (town, etc.)	J
Abstract	K
Deverbal	L
transitive	M
intransitive	N
capable of taking prefix NA	O
Adjectival	P
Nouns designating TIME	Q
days, year, etc.	R
events (birth, death, dinner, etc.)	S
Measure	T
preceded by a numeral	U
Inner state, emotion	V
Phenomena of nature (rain, frost, etc.)	W
Any inanimate, abstract or concrete noun	Y

Appendix II (cont.)

Verbs

Unspecified	X
Motion	1
Concrete action (chop, hit, tear, grind, etc.)	2
Motion in the broad sense of the term (swing, jump, bring, etc.)	3
Motion, intransitive and concrete action	4
Concrete state (sit, hang, stand, lie)	5
Concrete state and motion	6
Verb phrase	7
Inner state, feeling, intellectual activity (grieve, read, etc.)	8

Adjectives

Unspecified	X
Concrete (capable of being perceived and unmistakeably ascertained visually, e.g. tall, wide, white, etc.)	1
Abstract (denoting state, density, mood, age, temperature, sequence, sensed olfactorily or aurally, etc.)	2
Qualitative (good, bad, difficult, strong, etc.)	3

Appendix II (cont.)

Adjectives (cont.)

Denoting distance or location	4
Inner quality possessed only by humans or animals	5

APPENDIX III

Types of TRs abstracted in the initial stage of the study:

1. Spatial
2. Temporal
3. Causative
4. Purpose
5. Manner of action
6. Objective
7. General attributive
8. Similarity
9. Property
10. Togetherness
11. Possession
12. Condition
13. Adverbial attributive
14. Quantitative attributive
15. Designation
16. Attributive of absence
17. Origin
18. Spatial attributive
19. Separation attributive
20. Result-cause
21. Temporal attributive
22. Substitution
23. Restrictive attributive
24. Property attributive
25. Objective-restrictive
26. Restrictive in time
27. Similarity attributive
28. Designation attributive
29. Characteristic attributive
30. Container and contained attr.
31. Deprivation
32. Removal (ablation)
33. Conformity
34. Separation
35. Non-conformity
36. Comparative

Appendix III (cont.)

37. Class belonging

38. Detached object

39. Substitutability

40. Advocacy or defence

41. Transgressive

42. Distributive

The TRs are listed in order of their description in the Grammar.

APPENDIX IV

The 42 TRs were distributed into 11 major TRs as follows:

1. Spatial

2. Temporal

   Restrictive in time

3. Causative

4. Purpose

5. Objective

   Result-cause
   Advocacy or defence
   Transgressive
   Distributive
   Detached object
   Substitutability
   Substitution
   Objective-restrictive

6. Similarity

7. Togetherness

8. Deprivation

   Removal

9. Conformity

   Non-conformity

10. Comparative

11. Attributive

    Manner of action
    General attributive
    Possession
    Condition
    Adverbial attributive
    Quantitative attributive
    Designation
    Attributive of absence
    Origin
    Spatial attributive
    Separation attributive
    Temporal attributive
    Restrictive attributive
    Property attributive
    Similarity attributive
    Designation attributive
    Characteristic attributive
    Container and contained attributive
    Separation
    Restrictive
    Class belonging