Papers and Studies in Contrastive Linguistics 34, 1998 pp. 261-271

THE ON/IN ALTERNATION IN SERBO-CROATIAN AND POLISH: AN EXERCISE IN COGNITIVE LINGUISTICS AND DECISION THEORY

DANKO ŠIPKA
Adam Mickiewicz University

1. Introduction

This paper is aimed at answering if and how cognitive linguistics and decision theory can be useful in explaining a concrete phenomenon in two structurally, typologically and genetically closely related languages, in contrasting these two languages in this respect and, finally, in solving concrete problems involving the alternation mentioned in the title.

The prepositions have traditionally been a strong field of the cognitive linguistic research, and in any alternation one must decide which of the possible forms to use. This makes it a typical decision-making situation. This is why these two approaches have been chosen.

We assume the fundamentals of cognitive lingusitics and decision theory. The main propositions of cognitive linguistics can be found in Langacker (1987, 1990, 1991), while the core ideas of decision theory have been presented in works such as Keeney (1996) and Winterveldt and Edwards (1986).

2. Definitions and scope

The capitalized ON/IN in this paper are not English lexical items, but metalinguistic symbols the definition of which follows.

Under ON/IN alternation I understand the alternative usage of the prepositions *na* and *u/w(e)* in Slavic locative PPs depending on the properties of the landmark within the PP and its relationship toward the trajector outside it.

Since it should be clear from this definition that my starting point is the cognitive linguistic category of trajector-landmark construal, I will restrict myself to the simple atemporal relation represented by the mentioned Slavic preposition. In practice, this means that I will analyze locative PPs but not those in the accusative which represent a complex atemporal relation. Once the rules for the ON/IN alternation in

both languages are determined (as in case of the simple atemporal relation ON/IN + Locative) they can automatically be extended to the complex atemporal relation of the accusative case in Serbo-Croatian and to some extent in Polish:

na – ON, u, w(e) – IN, Ja sam, Jestem – I am, krov, dach (Nominative and Accusative), krovu, dachu (Locative), krova (Genitive) – roof, soba, pokój (Nominative), sobu, pokój (Accusative), sobe (Genitive) – room.

(1)

		na +	Loc	=	*na* +	Acc
S-Cr.	Ja sam	na	krovu.	Idem	na	krov.
Pl.	Jestem	na	dachu.	Idę	na	dach.
	(I) am	on	the roof$_{Loc}$	(I) go to		the roof$_{Acc=Nom}$
	I'm on the roof.			I'm going to the roof.		

		w(e)/u +	Loc	=	*w(e)/u* +	Acc
S-Cr.	Ja sam	u	sobi.	Idem	u	sobu.
Pl.	Jestem	w	pokoju.			
	(I) am	in	the room$_{Loc}$	(I) go to		the room$_{Acc}$
	I'm on the room.			I'm going to the room.		

and in Serbo-Croatian these rules can also account for the IZ/SA alternation:

(2)

		na +	Loc	=	*sa* +	Acc
S-Cr.	Ja sam	na	krovu.	Idem	sa	krova.
	I'm on the roof.			I'm going off the roof.		

		u +	Loc	=	*iz* +	Gen
S-Cr.	Ja sam	u	sobi.	Idem	iz	sobe.

There are further restrictions to this research. It takes into account only those locative PPs which pass the where-test, that is which answer the question "Where?" This provides only those locative PPs where landmark is a location in a very broad sense and excludes those where landmark is time, circumstance, etc. Finally, I am interested only in the general lexicon excluding any toponymic landmarks.

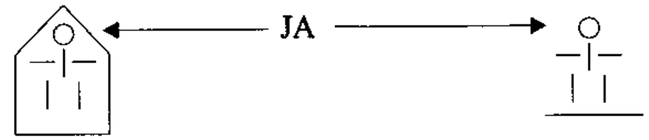
3. Starting Assumptions

In the trajector-landmark relation expressed by the locative PPs there are such landmarks which are construable in that they can be spatially represented ('room', 'house', 'tree', etc.) but there are also those which cannot be represented in that manner (e.g. 'meeting', 'war', 'elections', etc.). This is the first distinction relevant for the IN/ON alternation. It is namely so that landmark properties can be criterial for this alternation only in those cases where the landmark can be spatially represented. In that case the rule is as follows:

(3) *If landmark is mentally represented as an enclosed unit and trajector is inside this unit, use IN. Otherwise use ON.*

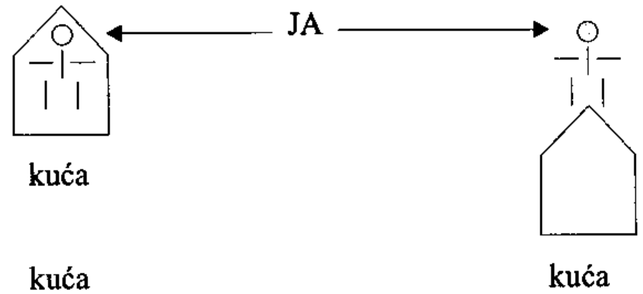
(4)

S-Cr.	Ja sam	u	kuci	vs.	Ja sam	na	plaži
	I am	in	house		I am	on	beach$_{LOC}$
	I am in the house.				I am on the beach.		



kuća plaža

s-cr.	Ja sam	u	kuci	vs.	Ja sam	na	kući
	I am	in	house		I am	on	beach
	I am	in	the house.		I am	on	the beach.



kuća kuća

This situation, where cognitive linguistics provides relevant explanation only in one segment of the alternation, brings us to the following questions this paper will attempt to answer.

Is there a difference in how native speakers process construable and non-construable landmarks?

- What is the approximate quantitative ratio between these two landmarks?
- Which decision-making criteria are to be used where landmark is a non-construable one?
- Which decision-making tree is the optimal one to model this alternation in both Serbo-Croatian and Polish?
- Is all this relevant in contrasting Serbo-Croatian and Polish?
- Is all this relevant in solving concrete problems involving these two languages?

In order to answer these questions, I have reached for psycholinguistic and corpus linguistic data.

4. Methodology

In order to discover if there are differences in processing construable and non-construable landmarks, a simple psycholinguistic experiment has been conducted where reaction time to these two landmarks were measured using two independent groups of subjects with group 0 representing construable landmarks and group 1 non-construable landmarks. The methodology of this experiment is summarized in Table 1.

Table 1. The methodology of the experiment

Design	Independent matched groups
Subjects	2x30 normal adult subjects, Polish native speakers, undergraduate students, freshmen and junior
Place	Adam Mickiewicz University, Poznań
Date	February 27, 1997
Type	Lexical decision task (10 sentences per group)
Independent Variable	0 – construable landmark, 1 – non-construable landmark
Dependent Variable	Reaction time (in hundredths of second)
Software Used	Answerometer (self-made), SPSS 5.0
Examples	000n Byłem __ pagórku. (w/n) 001n Byłem __ konsultacjach. (w/n)

If this distinction between construable and non-construable landmarks is relevant and if some fundamental claims within cognitive linguistics are true, we should expect higher reaction times to the non-construable landmarks.

To be able to answer the other questions stated above, I have also conducted a quantitative contrastive Serbo-Croatian – Polish corpus analysis using the texts described in Table 2.

Table 2. Texts used in the analyses.

	Polish	**Serbo-Croat**
Author	Dawid Warszawski	Milan Božić
Content	Articles about the war in the former Yugoslavia	Editorials about the war in the former Yugoslavia
Source	Polish daily *Gazeta Wyborcza*	Serbian radio *B 92*
Form	Electronic document	Electronic document
Period	9/21/93 – 12/7/95	6/3/93 – 2/4/95
Volume	979981characters 165566 tokens 42914 types	994043 characters 193890 tokens 52502 types
Language	Polish, standard, journalist	Serbo-Croatian, standard (Serbian variant), journalist

In both texts the following data have been collected:

- total frequency of *na* and *u/w(e)* in the corpus,
- frequency of those cases relevant to our research,
- frequency of construable and non-construable landmarks,
- frequency of *na* and *u/w(e)* with both construable and non-construable landmarks,
- frequency of the differences between Serbo-Croatian and Polish.

Finally, a qualitative analysis of these two corpora has been conducted in order to determine if a decision-making criterion can be found for the non-construable landmarks and if the proposed explanation for construable landmarks fits the explanation.

5. Results and discussion

The psycholinguistic experiment has revealed that there is a significant difference in processing construable and non-construable landmarks. Descriptive statistical data show that the subjects needed more time to process non-construable landmarks and inferential statistics show that this difference is meaningful. This can be seen in Table 3.

Table 3. The results of the experiment.

Table 3.1. Descriptive statistical analysis for construable landmarks (a) and non-construable landmarks (b).

(a) GROUP 0 (CONSTRUABLE LANDMARKS)

Variable	Mean	Std Dev	N
Time	193.82	112.81	300

(b) GROUP 2 (NON-CONSTRUABLE LANDMARKS)

Variable	Mean	Std Dev	N
Time	294.65	282.60	300

Table 3.2. Inferential statistical analysis.

(a) t-tests for independent samoles of GRP

Variable	Number of Cases	Mean	SD	SE of Mean
TIME				
GRP 0	300	193.8233	112.811	6.513
GRP 1	300	294.6533	282.601	16.316

Mean Difference = -100.8300
Levene's Test for Equality of Variances: F= 16.479 P= .000

(b) t-test for Equality of Means

Variances	t-value	df	2-Tail Sig	SE of Diff	95% CI for Diff
Equal	-5.74	598	,000	17.568	(-135.340; -66.320)
Unequal	-5.74	391.93	,000	17.568	(-135.377; -66.283)

- - - - - - - - - - - - - - - - - O N E W A Y - - - - - - - - - - - - - - - -

Variable TIME
By Variable GRP

(c) Analysis of Variance

Source	D.F.	Sum of Squares	Mean Squares	Ratio	Prob.
Between Groups	1	1525003.335	1525003.335	32.9412	.0000
Within Groups	598	27684263.58	46294.7552		
Total	599	29209266.92			

Table 4. The quantitative contrastive study of two corpora.

	Polish			**Serbo-Croat**		
	#	%		#	%	
total	8500	100%		7820	100%	
na	2827	33%		3017	39%	
w(e)/u	5673	67%		4803	61%	
			% (total)			*%(total)*
con.+non-con.	1739	100%	20%	2040	100%	26%
construable	834	48%		599	29%	
non-construable	905	52%		1441	71%	
difference	100	6%		105	5%	
construable	834	100%		599	100%	
na	377	45%		312	52%	
w(e)	457	55%		287	48%	
non-construable	905	100%	*difference*	1441	100%	*difference*
na	168	19%	1 (1%)	512	36%	104 (20%)
w(e)/u	737	81%	99 (13%)	929	64%	1 (0%)

The spatial IN/ON alternation in Locative PPs (as defined here) covers only a small portion (20-26%) of the distribution of these prepositions. But if we include this alternation in the accusative as well as toponyms, we can expect that it extends to more than one half of the whole.

It can be observed that Serbo-Croatian has a higher non-construable to construable ratio than Polish. This is possibly caused by the fact that their author was primarily interested in political speculation where the Polish author was more concerned with pragmatic, concrete events.

However, it is clear that this distinction is such that both possible options are open. It is not a situation where one is a default and the other an exception. Essentially, we have the 'either A or B', not 'unless B, A' relation.

It is interesting to note that the only discernible differences between Polish and Serbo-Croatian were found with non-construable landmarks – only a small percentage of the total distribution. Furthermore, in most cases, the Polish *w(e)* and the Serbo-Croatian *na*, helps explain the difference in the distribution of these prepositions in two corpora. These different forms belonged to 17 different lexemes.

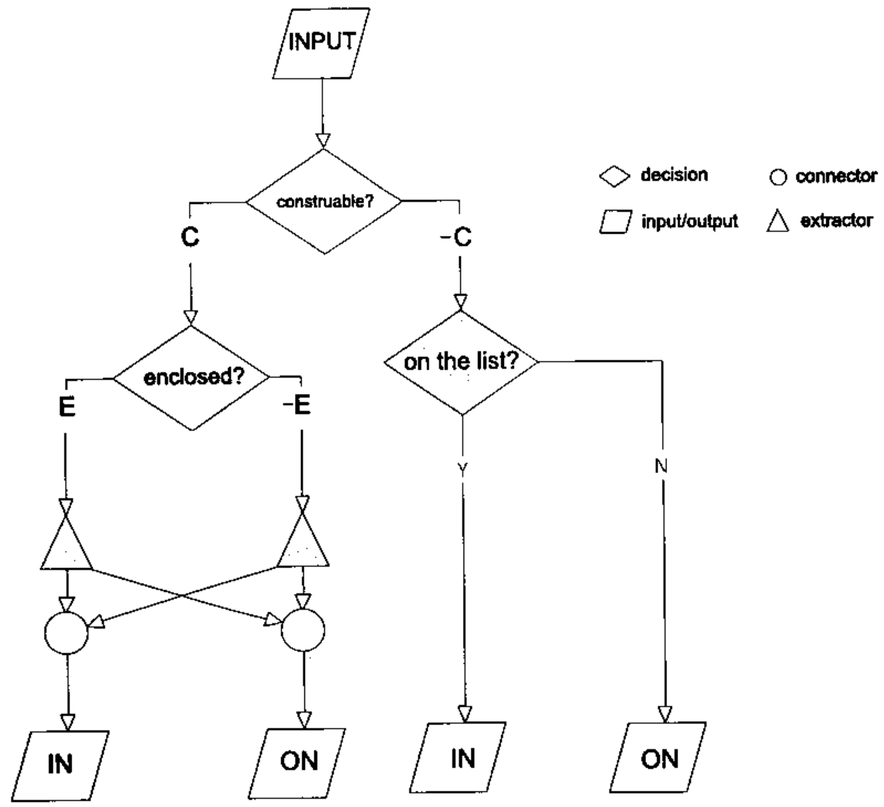
The qualitative analysis revealed that there is no single decision-making criterion as in the case of construable landmarks. The possible decision-making tree in this segment would contain a great number of nodes and exceptions. The alternative is to have only one node and a lengthy list of exceptions. A series of experiments is required to determine which of the solutions is more practical.

The qualitative analysis of Serbo-Croatian – Polish differences demonstrate that for some of these items one can find common hyperonyms, for example, electronic media (S-Cr. *na televiziji* – Pl. *w telewizji*, S-Cr. *na radiju* – Pl. *w radiu*), and politics (S-Cr. *na referendumu* – Pl. *w referendum*, S-Cr. *na izborima*, Pl. *w wyborach*).

6. A Possible Model

If we choose the one-node-with-list model for the non-construable landmarks, the decision-making flowchart (where we one must choose between IN and ON) can be presented as follows (the areas of Serbo-Croatian – Polish differences are shaded):

Fig. 1. The decision-making flow chart for IN and ON.



In those cases where the landmark is construable, the chart can account for practically all cases. The extractors are reserved for rare cases, primarily place names such as the Serbo-Croatian *Cetinje* or *Pale*, which take the *na* form rather than the *u* form.

7. Relevance of the Theories in Contrasting Serbo-Croatian and Polish

If we contrast these two languages using purely academic methods, the combination of cognitive linguistics and the decision-making theory provide two major observations, first, which segments are general and which are language specific. The distinction between construable and non-construable landmark is crucial in this respect. Furthermore, this distinction shows which segments are predictable and which are not. Secondly, the cognitive linguistic trajector-landmark relation serves as an almost perfect decision-making node to denote those cases where the landmark is construable.

However, because we normally only contrast two languages to facilitate concrete problem-solving tasks, these two theories (even when they complement each other) become almost irrelevant in this case.

Two typical problem-solving situations involving the IN/ON alternation of Polish and Serbo-Croatian are foreign language teaching and machine translation. For example, in foreign language teaching, it might be useful to warn students to exercise caution with non-construable landmarks. But the most efficient way of having them learn the difference between these two languages is to remember the list of differences starting with the most frequent cases.

In machine translation there are two typical solutions – either to compile a list or label lexical items so as to exhibit the difference. The label solution is used in MT program called NeuroTran (more information is available at http://www.tranexp.com) where, for example, the Polish – Serbo-Croatian lexical entries for prepositions appear as:

w LA/u LA
we LA/u LA
na LA/na LA

and this is how this preposition has been translated unless there is a label on the lexical item, such as:

telewizja,i f; w 1...236 3.4.2.12 .00323/televizija,e f; na 1...236 3.4.2.12 .00341

L – locative, A – accusative, 1...236 – usage sphere data, 3.4.2.12 - subject matter data, .00323 – frequency data.

Again, for this application, only a list of cases is required.

In this case, while cognitive linguistics has made an important step towards making the explanation simple and clear, when we merge it with another approach, it is practically of no use. More generally, I would maintain that there are situations

where one does not need a theory. This brings us to a number of serious questions, primarily those concerning the optimal ratio and relations between theoretic and applied research in linguistics which have yet to be answered.

REFERENCES

Keeney, R. L. 1996. *Value-focused thinking. A path to creative decision making*. Harvard: Harvard University Press.

Langacker, R.W. 1990. *Concept, image, symbol: The cognitive basis of grammar*. Berlin: Mouton de Gruyter.

Langacker, R.W. 1987. *Foundations of cognitive grammar. Volume 1: Theoretical prerequisites*. Stanford: Stanford University Press.

Langacker, R.W. 1991. *Foundations of cognitive grammar. Volume 2: Descriptive application*. Stanford: Stanford University Press.

von Winterveldt, D. and W. Edwards. 1986. *Decision analysis and behavioral research*. Cambridge: Cambridge University Press.