

first publication

This contribution was presented in  
Panel VI: (Inter-)regional Networks,  
December 11, 2003

Language, like society, consists of many different kinds of networks. Networks in a language depend on the societal, political and cognitive conditions under which they are employed. They do not occur naturally but are rather external and formed as a result of social interaction and practical usage. There exist various definitions of networks since they are understood and perceived in different ways. Thus I am not going to define what networks mean for me but I will try to give an insight and proposal on what they could be within this kind of approach in this paper.

In the first section, I will briefly introduce the Croatian linguistic situation. In the second section, I will provide some quantitative data and will try to establish some connections of linguistic items, which I will call networks. The last section will be a conclusion, including some remarks about the future of this kind of research.

### I. Introduction

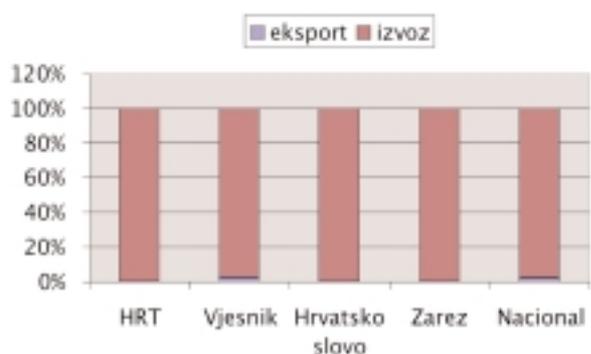
There are certain doubts about the existence of a Croatian language (not to mention Bosnian and Montenegrin) among linguists, sociolinguists, and other scholars dealing with South Slavic studies. The language we are talking about was only recently known as Serbo-Croatian alongside other commonly-known names. That is due to a genetic and linguistic similarity between Croatian and Serbian. Linguistically, there is no doubt that both idioms (languages) are very closely connected so that they represent one linguistic system (also called a diastem). Both standard languages are based on the same linguistic system (the neostokavian system), with some variations based on geographical location. However, their paths of development somewhat differed until Serbo-Croatian was officially and ideologically recognised as a language in the 19<sup>th</sup> century and was promoted as one during the 20<sup>th</sup> century in various periods and various strength of intensity. During the 19<sup>th</sup> century, Serbian was created in a revolutionary manner through the reform by Vuk Karadzic, while Croatian was standardized by evolution. Before 19<sup>th</sup> century nationalism emerged, it is difficult to define what the Croatian language was, since there were three main dialects that had had their own periods of development as standard, literary languages. Nowadays, we consider all of them, in spite of their linguistic differences to the new standard, as a part of a Croatian linguistic tradition that is represented in the lexicon by different sources (let us call them Mediterranean, Central European and Balkan). They created a huge corpus of linguistic items, notably lexemes, which did not belong to the new standard language but were a part of the tradition. Additionally, 19<sup>th</sup> century Croatian language purism created and coined more new lexemes, which survived in some literary pieces and notably in dictionaries (the most important one having been Sulek's dictionary). This lexical heritage is considered by the Croatian elite as truly Croatian and pure, in contrast to the Serbian language. Thus, in the periods of gaining political autonomy or independence, there are strengthening tendencies in language planning aimed at the revival of a »true« and »better« Croatian lexicon. This situation has created a lexical peculiarity. In terms of de Saussure's concept of *langue*, an abstract corpus of potential usage, there are plenty of lexemes, close synonyms, out of which one will be considered »better« and one »worse« or »less Croatian«. The main instrument of language planning (i.e. purist standardization in this case) are dictionaries and language guides in which the national ideology suggests using one, better, element and omitting the other, worse, one. We have to keep in mind that such purism can be very harmful for the language as every style or register (or discourse) has a different, socially constructed, corpus of appropriateness and formality or informality.

Standard languages should be understood as languages that are to some extent artificially constructed, created by centripetal forces that aim to establish and maintain something we are used to calling *nations*. Standard languages are cultural, political and most importantly *national* languages, particular choices made at a moment of national revival. In sociolinguistic literature there is a joking phrase that says that a language is a dialect within an army and a navy or that a standard language is a dialect that was lucky. From this perspective, every standard language is standardized in a very arbitrary way. The standardization that was carried out in Croatia through language planning aimed at purifying the language particularly in



The three charts and tables 1-3 prove that the analyzed pairs of lexemes are used differently in *HRT*, *Vjesnik* and *Hrvatsko slovo* than they are in the non-nationalistic media *Zarez* and *Nacional*. Nationalistic and state-owned media exclusively or overwhelmingly use the puristic equivalent while the non-nationalistic media use both lexemes simultaneously or display a clear preference for using the non-puristic lexeme. This gives us grounds to propose a hypothesis on a sort of ideologization of lexical usage taking place. Such a process, however, does not entirely equal a linguistic and notably semantic ideologization connected to a shift in the meaning of a particular lexeme/lexemes.

This sort of lexical usage is not the case in every researched example, however. We have found some data in our corpus that confirm that it is possible for one of the lexemes – either non-puristic or puristic – to dominate and be overwhelmingly used in all types of media.



<i>eksport</i>	0% (0)	2.3% (1)	0% (0)	0% (0)	2.1% (1)
<i>izvoz</i>	100% (36)	97.7% (42)	100% (4)	100% (2)	97.9% (46)

Chart 4

Chart 4 shows that only the purist lexeme *izvoz* is used and none of the newspapers use the internationalism *eksport*.

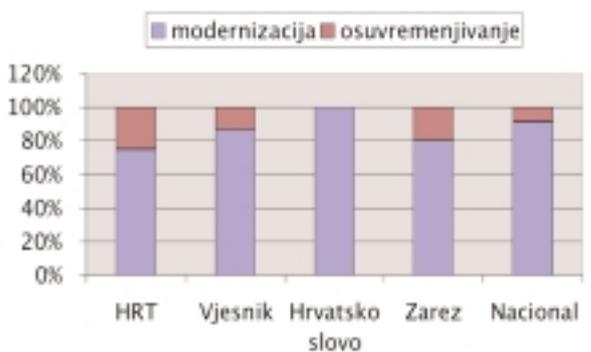


Chart 5

Chart 5, however, shows that only the internationalism *modernisation* is being used. We see that although the language-planning process towards this doublet has been under way throughout the 1990's, the lexeme preferred by the purist ideology could not be popularised in the analysed types of press.

## II.b Establishing national and non-national networks

As we can see, some newspapers more often use purist lexemes while some of them avoid them. We can assume that this is connected with their ideological attitude. We can conclude that we can expect a higher intensity of purist words in national and state-owned media as a realization of national ideology whereas in non-national media we expect a lower intensity as a manifestation of non-national ideology.

Let us group the whole corpus of 100 pairs presented here into clusters of five types of media. First of all, we have to establish quantitative frames for our corpus.

1. A pair in which a »better« lexeme is represented by a percentage of at least 60% has a *quantitative purist majority*.

<sup>2</sup> I am consciously using this term and I am not referring to its political and social meaning.

2. A pair in which a »worse« lexeme is represented by percentage of at least 60% has a *quantitative non-purist majority*.
3. A pair in which in turn both lexemes, the »better« one and the »worse« one, are represented by a balanced percentage i.e. not more than 59,9% each *does not have a quantitative majority*.

Now let us group all the pairs from the corpus of one hundred as they appear in the five analysed types of media: *Croatian Radio and Television, Vjesnik, Hrvatsko slovo, Zarez, Nacional*. The lexemes will be presented collectively in total numbers rather than representing each individual count.

*Hrvatski radio i televizija (HRT)* used the quantitative purist majority in 77 of the 100 cases, the quantitative non-purist majority in 10 of the cases and did not show a quantitative majority in 13 of the 100 pairs of lexemes.

*Vjesnik* used the quantitative purist majority in 69 cases, the quantitative non-purist majority in 16 cases and did not show a quantitative majority in 15 of the 100 pairs of lexemes.

*Hrvatsko slovo* employed the quantitative purist majority in 75 of the 100 cases and the quantitative non-purist majority in 15 cases, while it did not show a preference in 10 of the 100 cases.

*Zarez* used the quantitative purist majority in 41 out of 100 cases, the quantitative non-purist majority in 45 of the cases, and did not use either majority in 14 cases.

*Nacional* used the quantitative purist majority in 46 out of 100 cases, the quantitative non-purist majority in 35 of the cases and did not employ either lexeme majoritatively in 19 out of the 100 cases.

Chart 6 demonstrates these complete results and the frequency of lexeme usage in grouped pairs.

	<i>HRT</i>	<i>Vjesnik</i>	<i>Hr. slovo</i>	<i>Zarez</i>	<i>Nacional</i>	
quantitative purist majority	77%	69%	75%	41%	46%	
quantitative non-purist majority	13%	15%	10%	45%	35%	
no quantitative majority	10%	16%	15%	14%	19%	Chart 6

These findings give us some additional information and confirmation of the hypothesis given at the beginning of the paper. First of all, we can see that some newspapers realize a national ideology through its underlying linguistic purism more often than others. I grouped them into three clusters, as pointed out initially: the first is nationalistic (*Hrvatsko slovo*), the second state-owned (*HRT* and *Vjesnik*) and the third non-nationalistic and liberal.<sup>2</sup> We can see that state-owned media and national media tend to realize language planning more radically, i.e. the proportions of use of the officially preferred lexeme are 77% and 69% for the state-owned media and 75% for the national *Hrvatsko slovo* and 41% and 46% for the non-national press samples. On the other hand, the “worse” Croatian lexemes were used only in 13%, 15% and 10% of the cases in first two types of the media, while the non-national media used these in 45% and 35% of the cases.

Now we can make a comparison via mathematical calculation and specification of the whole set of the media. For each type of the media we subtract the number of “better” lexemes used from the number of “worse” lexemes and we will find the number that could indicate the degree of national radicalism and ideological conformism on the linguistic and systemic level for each of the media samples used. The higher the number we reach, the higher the national attitude that is represented in that media:

1. *Croatian Radio and Television*:  $77 - 13 = 64$
2. *Vjesnik*:  $69 - 15 = 55$
3. *Hrvatsko slovo*:  $75 - 10 = 65$
4. *Zarez*:  $41 - 45 = -4$
5. *Nacional*:  $46 - 35 = 11$

Comparing the numbers that were calculated, one can see how great the difference is between the types of the media I am analysing here:

*Hrvatsko slovo* (65) > *HRT* (64) > *Vjesnik* (55) > *Nacional* (11) > *Zarez* (-4)

As mentioned above, there is a certain number of lexemes from my corpus that occurs with a similar frequency in all the media samples. Additionally, there are six further pairs in the corpus that have been specialized semantically under the influence of language planning. I am, however, mostly concerned with doublets in which one lexeme has been adopted with at least 90% of usage. In the next calculation both former groups (frequent majority and semantic specialization) are not taken into consideration and I am using only the latter group, meaning that in chart 7, 54 pairs of lexemes are presented.

	<i>HRT</i>	<i>Vjesnik</i>	<i>Hr.slovo</i>	<i>Zarez</i>	<i>Nacional</i>	
quantitative purist majority	69%	55%	67%	11%	13%	
quantitative non-purist majority	11%	15%	5%	69%	54%	
no quantitative majority	20%	30%	28%	20%	33%	Chart 7

This chart allows us to carry out another calculation based on the same method as the previous calculation based on the results in chart 6, this time using the more selective results from chart 7:

1. *HRT*:  $69 - 11 = 58$
2. *Vjesnik*:  $55 - 15 = 40$
3. *Hrvatsko slovo*:  $67 - 5 = 62$
4. *Zarez*:  $11 - 69 = -58$
5. *Nacional*:  $13 - 54 = -41$

This calculation confirms our previous empirical observations. In terms of adoption of new, national lexemes, the most radical media sample used is *Hrvatsko slovo* and the least radical one is *Zarez*.

*Hrvatsko slovo* (62) > *HRT* (58) > *Vjesnik* (40) > *Nacional* (-41) > *Zarez* (-58)

This time, however, having excluding the stabilised pairs of lexemes and semantically specified terms, the numbers arrived at are even more convincing. The difference between the most national *Hrvatsko slovo* (62) and the most liberal *Zarez* (-58) widens dramatically to altogether a difference of 120 points.

### III. Conclusion

Since this presentation does not aim at analysing the sociolinguistic situation, I will limit myself to only one concluding observation: The difference between the most radical type of media (*Hrvatsko slovo*) and the less radical (*Zarez*) comes to 68 points [ $64 - (-4) = 68$ ] in the calculation with the first set of numbers and 120 [ $62 - (-58) = 120$ ] when calculating with the second set. This shows how diversified the realisation of language planning and thus the national ideology in the Croatian media landscape actually is. To sum up, I claim that the national ideology underlying language purism influences the lexical choice that is made. Thus, in estimating the national radicalism of the press we do not have to use textual analysis alone but can also rely on comparing the linguistic system. In conclusion, we can state that there are groups of networks of lexemes, which could serve as an indicator to discovering ideological networks in the Croatian media. Since this research is an extract from the bigger project of my doctoral thesis, the research for which was carried out in January 2002, it does not concern networks as such. However, my scholarly interests have been influenced only recently by the book *Linked. The New Science of Networks* by Albert-László Barabasi<sup>3</sup> which has led me to future plans to establish a whole set of such understood networks in the Croatian language and to widening my sociolinguistic research to a number of lexemes, which have so far not been included.

<sup>3</sup> Barabasi Albert-László, *Linked. The New Science of Networks*, Cambridge, MA, 2002

