

# Language, Policy and Identity: Perceptions of and Expectations for (Non)anglicized Language on the Web. The Case of Croatian Blogs<sup>1</sup>

**Marija BRALA**

University of Rijeka, Faculty of Philosophy, Department of English  
mbrala@ffri.hr

Questo articolo esplora il rapporto tra lingua, identità e politica linguistica, dal punto di vista dei blog creati in lingua croata. Le questioni esaminate includono: a) le tendenze linguistiche osservate nei blog croati; b) la percezione dei blog croati e di quelli 'anglicizzati'; c) le possibili implicazioni che le caratteristiche linguistiche osservate (e in particolare le preferenze degli autori di blog croati per l'inglese, ovvero per il lessico, la sintassi e le strategie di discorso anglicizzati) potrebbero avere per la politica linguistica in Croazia. Alla discussione generale del rapporto lingua-identità-politica nella parte introduttiva (sezione 1) segue una caratterizzazione delle blogosfere croate (2), un'analisi del linguaggio dei blog croati (3), un confronto tra i blog croatizzati e quelli anglicizzati (4) e, nella parte finale dell'articolo, una discussione delle osservazioni (5).

**Parole chiave:**

Lingua, identità, politica linguistica, blog

## 1. Introduction: on language, identity and policy

It is generally known and accepted that language plays a central role in the formation and expression of identity. The fact that the role of language and dialect in identity construction is becoming ever more central in the postmodern era, as other traditional markers of identity such as gender, race or class are being destabilized (see Warschauer, 2000), is less known – except by language policymakers, who apply it widely.

The language-identity-policy issue is multifaceted, viewable and interpretable from political, sociological, psychological, historical and many other perspectives. This work proposes a multidisciplinary approach to the problem, taking into account the sociological, or rather the sociolinguistic and psychological aspects of the intriguing link between the language in which a content is presented; the perception of that content; the public perception of the content's author (who creates in a particular language, or even in a particular linguistic subcode); and the possible implications these observations might have for those in charge of language policy. Departing from these general considerations, this paper focuses on two questions, 1) (how) does the language in

---

<sup>1</sup> I wish to thank two anonymous reviewers for their insightful and constructive comments on earlier versions of this paper.

which a content is presented influence perceptions of that content and of its creator(s), and 2) could and should these perceptions, which are based solely on language and thus constitute a language-identity feature, be related to the wider issue of language policy-making? Finding answers to these questions shall be separately dealt with in this paper.

Most people who have tried to communicate in two linguistic realities are familiar with the puzzling relationship between language (or rather, between bi- and multilingualism) and the perception of self<sup>2</sup>. Although discussions about bi- and multilingualism often posit that speaking different languages might mean feeling like and being perceived as a person with multiple selves, the topic has generally been viewed within scholarly circles as too naïve, simplistic, and heavily based on introspection for it to constitute a valid (psycho)linguistic variable. (Although, occasionally authors have taken to the study of multilingual selves, e.g., Grosjean, 1982). Only recently has a group of researchers (led by Dewale & Pavlenko, 2001) approached the study of bi- and multilingual selves with greater scholarly systematicity and vigour, trying to show that the issue's dismissal has been based on misleading, reductionist arguments, and that an in-depth study of the problem is of great relevance to a number of disciplines, from linguistics (especially SLA) to psychology (see Dewale, 2005; Pavlenko, 2006; Wilson, 2005).

While slowly finding its path into psycholinguistic studies, particularly those related to bilingualism, the issue of the relation between language and the perception of self – within the broader context of the language-identity issue – still seems largely neglected by sociolinguists. In particular, the issue has been neglected when studying implications of the interplay between language and the self on wider issues of language policy-making. This paper, among other goals, aims to point out this shortcoming and to address the need to reverse this trend.

Departing from the abovementioned work by Devale & Pavlenko (2001, 2003) and from their finding that speakers of multiple languages frequently perceive a linguistic and psychological split, this paper takes the claims about language and identity further. In this new study, data that have until now been purely self-perceptive evolve into identifying aspects of linguistic identity experiences that, while not directly observable in the study of individual identity perception, do become obvious within the realm of public discourse perception. The study thus explores the perception of Croatian blog sites on the basis of their linguistic features (mostly anglicized vs. non-anglicized traits). Its results,

---

<sup>2</sup> Furthermore, given the fact that most people are exposed to at least some degree of code switching (we do not use the same 'language' when we speak to family members, friends, or our superiors), one could even claim that the problem of language and identity common to most, if not all, language users. However, this issue is outside the scope of this paper.

aimed at providing elements relative to the language-identity binomial, also provide useful insights for those seeking to identify and understand the sources of (self-)perceptions. This group should, among others, necessarily include language policymakers, since an understanding of psycho- and sociolinguistic aspects of language perception phenomena is needed for the adequate analysis, creation, and implementation of language policies.

One must mention that in the case of Croatia, the language policy issue is a particularly sensitive one. For the past fifteen years, following the civil war<sup>3</sup>, language policy in Croatia has had a clear mandate: Promote or prescribe strong national linguistic choices in all fields, frequently regardless of usage. This has at times meant a return to archaic, forgotten lexical and, occasionally, syntactic options, as well as leading to a systematic replacement of English or anglicized lexemes with Croatian translations. The driving force behind this mandate was national-political. In the 1990s, a main goal of language policy in Croatia (which, given the confusion, or rather the inexistence of a single language policy-making body or person<sup>4</sup>, usually boiled down to mere political talk about language) was to discourage the use of linguistic elements that were shared among the linguistic variants of former Yugoslavia, and foremost to widen the gap between Croatian and Serbian, two languages that until the 1990s were treated as one, referred to as Serbo-Croatian or Croatian-Serbian<sup>5</sup>. As we shall see in Section 5, such a policy has its problems. Among other things, it can result in a negative trend, namely in a rejection of the 'forceful anti-internationalization' mandate that leaves insufficient room for the linguistic reality of usage. In order to resolve the issue, language policy creators must constantly work to find the right balance between the set norm and the actual usage, but they must also try to understand those psychosocial elements that shape the perception of linguistic choices made by language

---

<sup>3</sup> Croatia, formerly one of the six republics that formed Yugoslavia, was recognized as an independent country in 1992, but the war for independence did not end until 1995.

<sup>4</sup> The question of who the language policymakers in Croatia were has no straightforward answer. Croatia has no central governmental or institutionalized body that acts as the official language policy creator. The body that comes closest to having this mandate is the 'Vijeće za normu hrvatskoga standardnog jezika' ('Council for the norm of the Croatian standard language' of the Institute for the Croatian language and linguistics, see <http://www.ihjj.hr/#vijecezanormu>), but this is a purely advisory body. This means that in practice different institutions, bodies and individuals have taken it upon themselves to 'prescribe' the norm, resulting in some contrasting norms. Given all the above, this paper does not criticize Croatian language policy, but rather the lack of a clear policy, especially one that would consider issues such as the relationship between language and identity.

<sup>5</sup> In order to complete and clarify the picture, it is extremely important to know that 'Serbo-Croatian' was also a product of an ideology, namely that of the unification of peoples, i.e. of the nations contained within the boundaries of former Yugoslavia. In practice, however, these peoples spoke either the Croatian or Serbian variant of 'Serbo-Croatian', even during former Yugoslav nationhood.

users. It is important to note here that part of the problem resides in the fact that critics of internationalisms in Croatian have one general approach to '(the) language as such'. Instead, what might be required is a certain degree of flexibility, for practice has shown that some jargons (Internet, fashion, music industry, *et al.*) tend more toward anglicization than others, and these linguistic subcodes need not necessarily be subject to the same criteria as standard language. The issue of norm vs. usage, as well as the issues of linguistic preferences and perceptions, are addressed below and examined with respect to the language of Croatian blogs.

After a brief introduction to the issue of blogs and blogging in general, the paper analyses the typical features of the language of Croatian blogs and then moves to present the results of a study aimed at investigating the perception of blog sites, contents and authors based solely on linguistic features. The results are followed by a discussion, in turn followed by a conclusion and suggestions for potential future research directions.

## **2. Blogs, bloggers and blogging**

In this section we explore those features that seem to characterise blogs, the typical behaviours and expectations of bloggers, and we then analyse the traits that characterise the behaviour termed as 'blogging'. Particular attention is paid to those features that differentiate the (linguistic) behaviour in question from other (computer-mediated) modes of communication.

### *2.1 About blogs in general*

The language used on the Internet for a number of reasons. For one, it presents a very rich arena for investigating the evolution of discourse (Crystal, 2006). A particularly interesting subtype of Internet language or *netspeak* (see below) is the jargon used by young people (Greenfield & Subrahmanyam, 2003). It is characterised by a combination of adapted traditional forms and new written forms, which include adaptations of slang, acronyms, and other non-standard forms of various provenance, newly coined forms, cross-linguistic forms, foreign forms (both in terms of lexis and syntax), paralinguistic elements (e.g., emoticons), and others.

Although scholars have actively explored the traits that characterise language in on-line interactions (such as chat rooms or newsgroups), the astonishingly fast evolutionary pace of both hardware and software development poses a great challenge to everyone interested in computer-mediated communication (CMC). For with practically each new Internet application that is created and embraced, a new type of discourse emerges. Blogs represents one excellent, illustrative example of the wide-ranging implications stemming from the rapid creation of discourse types.

Blogs are personal Web journals or reverse-chronological commentaries, written by individuals and made publicly accessible, a type of Web application that allows users to create, enter, display, or edit their posts at any time, as well as to comment on other blogs (cf. Crystal, 2006: 240). Immediately noticeable is the blog's private (content) vs. public (posting) character, which, as we shall see in more detail below, is of particular interest for linguistic analysis.

Blogs have distinctive technological features that set them apart from other forms of CMC (Herring *et al.*, 2004; Huffaker, 2004; Huffaker & Calvert, 2005), including: 1) ease of use, as users do not need to know HTML or other Web programming languages to publish on the Internet; 2) free access; 3) simple ways to archive information and knowledge; 4) opportunities for others to comment or provide feedback for each blog post; 5) links to other bloggers to form on-line communities; and 6) links to other media (sound, video, etc.). These technological features that characterise blogs translate in the user's world into a particular subtype of netspeak, which, given its clear distinctiveness, we shall call *blogspeak* (to be analysed in Section 3).

As already mentioned, blogs are a prolific field for (socio)linguistic analysis. The technological features described above have contributed to their popularity, and although it is notoriously difficult to gather statistics of blogging (Crystal, 2006: 246), even rough estimates relative to the number of active blogs in the past few years are impressive and reveal the size and impact of the phenomenon:

Estimates of the number of active blogs in 2003 varied from 1 million to 3 million. Commentators were suggesting that the number of blogs was doubling every six months. Estimates in 2005 were typically in the region of 12 to 15 million, but some were as high as 50 million worldwide. In April 2005, Microsoft reported that over 4.5 million 'spaces' (blogs) had been created since January of that year. Updates were running at around 170.000 a day. Of course, by no means all these new sites remain active [...] On the other hand, a popular blog attracts a huge readership, often over a million a month, which is more than the audience reached by many newspapers. The French music station, Skyblog, was reporting 1.6 million users in 2005 – 2.6 percent of the population of France. (Crystal, 2006: 246-247)

If to these data we add the 2007 figure<sup>6</sup> that reveals that currently there are over 70 million active weblogs, and that about 120.000 new weblogs are being created worldwide each day (which also means that 1.4 blogs are created every second of every day), it is easy to see the size and importance of the phenomenon. In slightly different terms, it is clear why 'blog' was already in 2004 declared 'word of the year' by dictionary publishers Merriam-Webster.

Given all of the above, the language of the blog – as a particular subtype of netspeak – is beyond doubt becoming an increasingly influential vehicle of

---

<sup>6</sup> For these statistics see <http://dijest.com/bc/>.

communication, likely to impact society also in fields that are not necessarily Internet-related, and thus deserving of scholarly attention.

The next section analyses blogsppeak within the Croatian blogosphere, i.e., blogs written by authors in Croatia in the Croatian language (or rather, it analyses its substandard version, namely Croatian blogsppeak) and determines which perceptions and expectations the public has with respect to blogs, based solely on the language, the sub- and paralinguistic codes used.

## 2.2 *Into the Croatian blogosphere*

As the saying goes, no one on the Internet knows if you're a dog, or if you're black or white, male or female, gay or straight, rich or poor. But people can immediately notice what language and dialect someone is using. The main question underpinning the studies reported in this paper is, What conclusions can be reached relative to a blog and its author, based solely on language? We first conducted a straightforward analysis of the features that characterise the language of Croatian blogs. Next, we created two blogs that differed in language (one in standard and the other in blogsppeak) and observed the hits and comments on these two blogs. The aim was to contrast the blog content perceptions and expectations, based solely on the contrasting 'anglicized' vs. 'croaticized' language. Finally, we undertook an elicitation response study aimed at investigating the perception of blogs and bloggers by a group of university students.

## 3. **Study 1 – the language of the Croatian blogs**

### 3.1 *Methodology*

In order to examine the language used in the Croatian blogosphere, we conducted a linguistic form analysis of randomly selected blogs. These blogs resided on the two most popular Croatian blog hosting sites that provide free hosting and administration services, namely blog.hr (n=30) and blogger.hr (n=30). These two sites host the largest number of Croatian blogs, are both very simple to use and open without limitations to the public. Both hosts offer various categories that classify blogs by primary content areas such as 'entertainment blogs', 'writing blogs', or 'teen blogs', and our sample contained blogs from all categories. Although different categories might be expected to be characterised by different (sub)linguistic varieties – e.g., a teen blog would be expected to differ from a political blog – a preliminary analysis conducted prior to this study has shown that the language and discourse patterns in all blog categories share similar, typical traits.

We intentionally excluded from our analysis the blogs by Croatian politicians, journalists, and other public personae who also post on the two blog hosts

examined. Preliminary analysis has shown that the language used in blogs authored by public names (or by whoever writes on their behalf) differs qualitatively from that used by 'nicks', anonymous writers who post and sign their blogs with a nickname. The former use blogs more as advertising spaces, and their language is much more formal both in style and content, much closer to the standard than the language of the latter group. Due to lack of space, this issue is not dealt with in detail, but it seems to be a point deserving further analysis.

### 3.2 Procedure

The front page of each weblog was analyzed, including its comments, and the most common characteristics were grouped into descriptive categories. We examined the overall characteristics of the discourse as expressed in language and paralinguistics, paying particular attention to non-standard forms of both syntax and lexis. Anglicized forms constituted one particular subtype of non-standard semantic and syntactic varieties from the sample.

### 3.3 Participants

Participants were randomly selected weblogs, for a total of 60 (30 blogs per blog host). It is difficult to ascertain the identity of weblog creators, as most authors create their blogs using a nickname, so the information given on the blog cannot be verified. However, some studies help identify the typical blogger.

Regarding the blog population globally, Huffaker & Calvert (2005) find that most surveys suggest that a significant portion of the total blog population is teenaged and also cite the *Perseus Development Data*. This data finds blogs dominated by youths, with 52% of all blogs being developed and maintained by teens aged 13-19, and with the gender split of all bloggers being 56% female and 44% male.

The situation of the Croatian blogosphere differs slightly. It has been suggested (Japundzic, 2006) that 88% of bloggers (which includes 'active' bloggers, blog authors and blog commentators) in Croatia are between 20 and 40 years of age, and 52% are between 20 and 30. The largest group of bloggers is people between the ages of 26 and 30, who total 30% of bloggers<sup>7</sup>.

---

<sup>7</sup> The age group analysis of the blogosphere in Croatia is as follows (after Japundzic, 2006): up to 15 years of age - 3%; 16 to 20 years of age - 9%; 21 to 25 - 22%; 26 to 30 - 30%; 31 to 35 - 11%; 36 to 45 - 25%; 46 to 55 - 0%; older than 56 - 0%. We do not see an explanation why bloggers in Croatia seem to be older than those in other countries. As the methodologies used to determine the mean age of bloggers in the two studies (*Perseus Development* vs. Japundzic) are not identical, this might explain at least part of the age gap. Other studies reported monthly by various Croatian Internet sites suggest a larger portion of teens in the general blogging

### 3.4 Results

The results are grouped into three categories, relative to a) lexicon, b) syntax and discourse, and c) other (paralinguistic) features.

#### a) *Broken, substandard lexical forms*

- *Specific jargon*

This includes Internet-related words, or more precisely, words related to the blogosphere (such as 'post', 'blog', 'blogger', 'blogerica', 'coolest link'). These words all retain the English form, with occasional spelling and morphological adaptations, or rather, croatizations – such as *bloger* instead of 'blogger', or 'blogerica' (created from 'blog' + 'er' [infix meaning 'person doing thing'] + 'ica' [ending for feminine]).

This feature category also includes newly coined words used exclusively by young bloggers, the meanings of which are not transparent. Examples include phrases such as 'Osjećam se jazzie' ('I feel jazzie' – note the anglicized 'jazzie', where the Croatian phonetical spelling would be 'džezi').

- *Distinctive, substandard orthography*

Concerning the differences between standard Croatian and blogspeak Croatian, the orthographic feature that most obviously varies is the use of capital letters and punctuation. Much blogspeak, like much Internet language, is not case sensitive, and there is a strong tendency to use lower case. This also means that capital letters and words in capital case are strongly marked (e.g., 'shouting', marked as very important, or other types of extra emphasis – cf. also Crystal, 2006: 92).

- *Spelling variations*

Spelling variations include all departures from the standard spelling system. There are two main subgroups within this typical trait of blogspeak. The first group, a very frequent one, includes Croatian (shallow) orthography of English words. Illustrative examples include *onlajn* for 'on-line', *ril lajf* for 'real life', *kul bener* for 'cool banner', *frend* for 'friend', *bič* for 'bitch', *partibrejker* for 'party breaker', etc.

Another group includes anglicized spelling (in part or entirely) of Croatian words, such as *maximalno* instead of 'maksimalno'.

Within the aforementioned group, we find a particular subgroup where English and Croatian words (and spelling) are mixed. A good example of this is the signing form *kiss swima* (a curious formula to say 'a kiss to everyone'. Note the

---

population than Japundžić's (ibid.) study. However, these cannot be taken as entirely reliable sources, as their sampling methodology is not made explicit).

'w' in 'swima', which should actually be spelled 'svima', meaning 'to everyone', plus the combination of this word with the English 'kiss', which has the translational equivalent of 'poljubac').

Both latter groups offer an interesting peculiarity. Even though (or exactly because) the Croatian alphabet does not have the letters 'x', 'y', or 'w', we find the 'ks' frequently replaced by 'x' (as in *paradoxalno* instead of 'paradoksalno'), the 'j' or 'i' replaced by 'y' (this is particularly frequent with people's names, such as Kety instead of 'Keti', or *Amelya* instead of 'Amelija'), and 'u' and 'v' replaced by 'w' (*swima* instead of 'svima'). This tendency shows a clear preference for the anglicized form, which is seen as 'more international', 'more modern', 'fuller', 'more elegant', 'more interesting' (see Section 4 below).

- *Abbreviations and acronyms*

Croatian blogspeak is further characterised both by abbreviations of Croatian words (e.g., 'stud' for students, 'prof' for 'profesor', 'komp' for 'kompjutor') as well as – and much more so than in the case of Croatian abbreviations and acronyms – by the use of English acronyms. Out of the list of abbreviations used in netspeak conversations proposed by Crystal (2006: 91-92), the 60 blogs analyzed contain the following abbreviations, used in both upper and lower case forms: **asap** (as soon as possible); **btw** (by the way); **cu** (see you); **gr8** (great); and **thx** (thanks). Not included in Crystal's list but found in Croatian blogs were: **ILU** (I love you); **ILU2** (I love you too); **4U** (for you); and **Xx** (kisses).

- b) *Broken, substandard syntax and discourse patterns*

- *Anglicization of NP patterns*

A typical Croatian noun phrase (NP) structure is a noun (N) premodified by an adjective. However, frequency of use suggests that this NP structure is increasingly abandoned in favour of the anglicized noun, premodified by noun structure (e.g., we find 'Internet stranica', a calque of 'Internet page', which according to standard Croatian grammar should be 'Internetska stranica', namely 'Internet's page'; or 'chat room' ('chat soba'), which is used instead of the standard 'soba za cavljanje', or 'room for chatting'). An acceptability rating study by Starcevic & Geld (2007) has shown that between the anglicized noun + noun construction vs. the proper Croatian adjective modifying the noun (NP) construction, the larger public favours the former. This is a clear sign that Internet jargon features (since most N+N constructions in Croatia come from the Internet and marketing jargons) quickly spread into everyday language.

- *Sentence types and (new) discourse strategies*

The syntax of Croatian blogspeak varies widely, ranging from very simple, almost telegraphic sentences, to complex sentences with a series of embeddings. Most interesting, these features are not homogenous in terms of

being mutually exclusive, but the combining of different syntactic patterns also results in substandard discourse types. All traditional markings of discourse cohesion and coherence can be broken, and the new discourse pattern resembles a stream-of-thought model. This feature once again reminds us that blogs are written in a highly intimate fashion, where one is not obliged to observe the rules and regularities of language (as if one were writing for oneself). This presupposition of intimacy, however, is obviously false: Blogs are posted on the most public medium in the world, the Internet. The clash between the private (strategies) and the public (use or rather reception) with respect to blogging deserves further attention, also from the viewpoint of discourse analysis.

c) *Paralinguistic features*

▪ *Use of emoticons*

Blogs are graphical texts; emoticons are perhaps the most typical traits of the sample of blogs in our examination. Emoticons are used as standard and well-established text elements in a variety of forms. It is our impression that emoticons are used more in blogs than in any other type of Internet communication (emails, chat groups, etc.). Most frequently used emoticons include:

:) , :))))) , animated emoticons, :( , :-P , :-PPPP, and =) .

It should be noted in this context that emoticons are far more frequently used in commentaries than in blog posts.

▪ *Multimediality*

Blogs are also multimedial sites. Apart from texts, they almost without exception contain links to video and audio material (the favourite being YouTube). This also leads to a frequent shift between languages, most typically between Croatian and English.

Three questions arise at this point:

- What are the motives, or rather the psycholinguistic, sociolinguistic, and other forces that drive the linguistic developments observed in 2.2? In other words, in which way does the specific Internet jargon reflect and create the identity of its users, and, at a more general level, what is the role of language in identity construction?
- What are the wider linguistic implications of the linguistic developments observed in 2.2, or in other words, could blogging or Internet use in general be seen as a phenomenon determining at least a degree of language shift toward English (as a lingua franca of the Web), not only as pertains to specific 'e-jargon', but also in relation to the standard?

- Are there conclusions to be drawn from the above analysis that could be useful to language policymakers? In other words, is it possible and necessary to establish a language-identity-policy triangular relation?

Although these three questions are complex, and although it is impossible to address, let alone satisfactorily explore and answer them in one scholarly paper, the remainder of this article does attempt to pinpoint at least some causes and implications of new linguistic developments stemming from technological innovations, primarily those related to computer mediated communication. In order to do so, we conducted a study aimed at researching the issue of language and identity of the blogger. Because our analysis of Croatian blogspeak revealed a strong influence of the English language, our study focused on whether there existed a difference in the perception of anglicized vs. non-anglicized blog content, and if so, what characterized such a difference in perception.

In order to investigate this issue, we conducted the following study.

## **4. Study 2 – anglicized vs. croaticized blogs**

### *4.1 Methodology*

The study was conducted in two phases. In the first phase, two blogs, one on [www.blog.hr](http://www.blog.hr) and another on [www.blogger.hr](http://www.blogger.hr), were created. The two blogs were identical in content: a 10-day–10-entries Web diary of a general, common nature, containing occasional comments on the language of the Web, and a link to the comments section. However, the blogs differed in the language in which they were created.

One blog (host [blog.hr](http://blog.hr)) was created under the nickname 'sestra-j' (Croatian for 'sister-j'), and the name of the blog was 'blogopitnik' (a compound of *blog* + *upitnik*, the Croatian word for 'questionnaire'). The blog was written entirely in standard Croatian and contained no emoticons, other paralinguistic forms, or anglicized forms. (The Croatian equivalent was used whenever possible instead of an anglicized form.)

The second blog ([blogger.hr](http://blogger.hr)) was created under the nickname 'sister-jay', and the name of the blog was 'blogonnaire' (a compound of *blog* + *questionnaire*). With this blog, the language used had all characteristic features outlined in 3. above, that is, it made ample use of emoticons, of anglicized semantic and syntactic forms, and of blog jargon.

Both posts remained active for a month (May 1<sup>st</sup> to May 30<sup>th</sup>). We counted the number of visits to each post, as well as counted and analysed the comments made on each of the posts.

It should be pointed out that any conclusions based solely on the number of hits on a site can be challenged. First of all, as pointed out by Wright (2004), accessing a site does not necessarily imply reading its contents. Further (although our preliminary evaluation of the Croatian blogosphere seemed to exclude this possibility), numerous parameters outside our control might influence the amount of access to blogs on one host compared to the other.

Given the above, we prepared phase two of the study, which consisted of an alternative research strategy to gauge blog attractiveness, a strategy of blog evaluation. The methodology selected was response elicitation. In this second phase, 38 university students were first given a questionnaire containing the following questions:

1. You have to choose one of two blog links: 'www.blogopitnik.hr' (link A) or 'www.blogonaire.hr' (link B). Which one do you choose?
2. The two blogs offer a series of links (see below).
  - A) Fun / nightlife / chat / wallpapers / e-cards
  - B) Zabava / noćni život / pozadine / el. razglednice<sup>8</sup>
 Which links do you choose?
3. If you write / were to write a blog about everyday life, which nickname do / would you choose?

Second, having completed this part of the questionnaire, study participants were given printouts of the title pages for the 'www.blogopitnik.hr' and 'www.blogonaire.hr' blogs, then were asked to rate the two blogs in the following categories:

4. Blog more likely to be up to date
5. Blog more likely to be interesting
6. Blog more likely to be useful
7. Which of the two blogs is clearer (clear language)
8. Blog more likely to be placed among my favourites

The ninth, final question was open ended:

9. Could you please make a comment relative to what motivated your preferences (if you had any).

All questions were posed in the Croatian language<sup>9</sup>. The answers given by students and reported below were translated into English by the author

---

<sup>8</sup> These are Croatian translation equivalents for the categories under 2 A).

<sup>9</sup> This is important, since, as shown by Brala (2007), when it comes to researching issues of language (bilingualism) and identity, there is a strong correlation between the results and the

of this paper, and translations were verified with an English native speaker who understands Croatian.

## 4.2 *Participants*

While, as already pointed out, it is impossible to ascertain the exact identity (age, gender, educational background, etc.) for the participants in phase one, and while all observations relative to bloggers in Croatian outlined in 3.1 apply in this case as well, we do have the data of all participants who took part in the second phase of the study. The subjects were 38 university students from the faculty of philosophy of the University of Rijeka in their second, third, or fourth year, who study a variety of subjects (English, Croatian, history, art history, information technology, among others). Most subjects (73%) rate themselves as Croatian-English bilinguals, some of them (26%) speak a third language (usually Italian or German), and in three cases even a fourth. Twenty-six subjects rate themselves as occasional blog readers, and 12 subjects rate themselves as regular blog readers. Nine subjects (out of the 12 regular blog users) are also authors of blogs.

We are fully aware that the subject sample chosen for our study is homogeneous in terms of age, education, linguistic picture, social background, and that as such it does not represent an ideal sample for generalization. However, we wanted to discover the perceptions and preferences of subjects who (following Wright, 2004) are a) among the most computer literate in the society, b) have regular (free) access to computers, and c) are at a level of study at which they are asked to do independent research. This means that our subjects have the skills, the opportunity, and the need to use computers and the Internet, and do as such represent an ideal group for a study on Internet usage practices. Furthermore, linguistic preferences or practices observed among university students are likely to point to at least some of the *developing practices* within society.

It should be mentioned that simultaneously with the study reported here, we attempted to conduct a parallel control study among the population aged 50-60, who are not regular Internet or blog users. The first results of this study proved relatively inconclusive, as most respondents aborted the completion of the questionnaire, explaining that they did not know enough about the subject matter and thus could not give pertinent answers. We thus do not include this group in this final report on the study.

---

language in which the subjects were tested or surveyed. Because subjects in our case were tested in Croatian, we can infer that the preferences expressed for anglicized forms might have been even greater if the survey language had been English (for details see Brala, 2007: 74-77).

### 4.3 Results

For the first part of Study 2, the numbers of 'hits' or visits to the two blogs (blogopitnik.hr and blogonnaire.hr) registered in May 2007 and the number of comments posted by blog visitors for each of the two blogs over the same period are shown in Fig. 1 below.

Blog	N° of visits from May 1 <sup>st</sup> to May 30 <sup>th</sup> 2007	N° of comments from May 1 <sup>st</sup> to May 30 <sup>th</sup> 2007
Blogopitnik.hr	17	0
Blogonnaire.hr	42	17

Fig. 1: Statistics of visits and comments for blogs

Two things related to comments must be noted at this point. First, there were 17 comments on the 'blogonnaire' blog and none on the 'blogopitnik' blog, a result that might have other causes, such as technical problems with the comments section for 'blogopitnik'. Unfortunately, this could not yet be ruled out as a cause at the time this article was written. Of course, other possible explanations for the result exist, such as a different demographic of individuals accessing the site, less 'appeal' of the blogopitnik than the blogonnaire site, and so forth; the latter reasons might at least in part explain the difference in the number of hits on the two posts. However, these explanations remain speculative, and investigating them fully is beyond the scope of this paper.

Second, the 17 comments posted on 'blogonnaire.hr' were all highly supportive of English terms, that is, of anglicized/internationalized Croatian language for the Web. Reasons proposed in support of this position can be grouped into four categories:

- a) the Web is global, so its language should be, too; the 'non-standard' Croatian is actually the 'standard Croatian of the Web';
- b) problems with Croatian diacritic signs; only English keyboard or software;
- c) by limiting language, we limit audience; the Web is free and open, so Web language should be the same;
- d) English/international variants are more likely to come up in search engines.

We now move to the results of phase two of the study, to the answers given in the questionnaire administered to university students.

Answers to **Question 1** (you can read only one blog; between 'blogopitnik.hr' and 'blogonnaire.hr', which one do you choose) and **Question 2** (do you click on links in English or in Croatian) are presented in Fig. 2 below.

Question	Anglicized version (blogonnaire)	Croatianized version (blogopitnik)	Don't know / did not answer
N° 1 ('blogopitnik' or 'blogonnaire')	26	5	7
N° 2 (anglicized or Croatianized links)	24	8	6

Fig. 2: Responses to Questions 1 and 2 of the questionnaire

The responses to **Question 3** (choose a nickname for your blog) follow:

Foreign/anglicized names: Lola; Sweet Angel; Mersault; Cherry; Sue-Chiara; Natalya; Ocean Wawe; Horse whisperer; Dark Angel; Sweet 666; Rinaa (shortened form of the name); Leeloo – my standard Net Nick; No angel; Tweety; Hot-chocolate; Shark; Anchoress; Extatic Epicurean; John Doe; Hovercraft; Merlin; Betty Boop; Wild Rose; Abutterfly; Smiley; Snowflake; Tinkerbell; Tammy, CroGirl; Roaring wind, Blue, Little Baby, Philya - TOTAL 33

Croatian proper names: Tanja, Rina – TOTAL 2

Croatian nicknames: Traktorica, nebuloza – TOTAL 2

Other names: ich – TOTAL 1<sup>10</sup>

The answers to **Questions 4 through 8** are presented below:

Question	Anglicized version (blogonnaire)	Croatianized version (blogopitnik)	Don't know / did not answer
N° 4 (which is more up to date)	27	5	6
N° 5 (which is more interesting)	22	6	10
N° 6 (which is more useful)	14	4	24
N° 7 (which is clearer)	33	2	3
N° 8 (more likely to be 'favourite')	30	4	4

Fig 3: Responses to Questions 4 through 8

The answers to **Question 9** are grouped below. As evident, not all individual answers are listed; rather, we have grouped replies into categories based on

<sup>10</sup> By a student who has German as B major.

sameness or close similarity of argument and thought. Groups are ordered in terms of frequency, from highest to lowest:

- The anglicized version of the name, i.e. the anglicisms in blog/Internet language are more common, we are more used to them, they sound better; they 'feel' more appropriate, more adequate, more 'exact' (proper jargon). Vice versa, the 'croaticized blogs' language seems sloppier, less elegant, less 'exact', less informed, a bit 'archaic'.
- I always have problems understanding the Croatian translations of English Internet jargon. English is definitely clearer, also in the Croatian context. Croatian translations of English Web terms are unclear, they seem like neologisms, like 'ideological' choices that someone is trying to force upon us. It is just not the 'natural' language of the Web.
- Whenever I see Croatian words instead of the usual English terms related to the Web, I get the feeling that the person writing is an amateur, does not know enough about cyberspace. I always wonder whether a person who uses Croatian words instead of the usual English ones follows current affairs at all.
- I cannot explain why, but I am naturally drawn to the English, i.e. the anglicized language; I am drawn to the English version intuitively, subconsciously, cannot explain why.
- All croaticized versions of English Web terms/contents are longer, more complex and difficult to read. When I see such a text, I generally abort without even attempting to read. It simply does not appeal.
- Croatian Web language is senseless! English inspires! English words are richer, more meaningful; Croatian sounds 'pseudo-scholarly'; Croatian sounds too formal and not immediate.
- To me, all the best on the Web is linked to the English language, perhaps because my first encounter with the Web was through an English site, plus I use the English version of Microsoft Office, it is partly a habit, but partly also a need. Why confine oneself?
- English is as 'open to the world' as the Internet. English is as international as the Internet. Croatian on the Net is total nonsense, it is contrary to the (global) logic of the medium; Croatian translations do not have the same meaning as the English source terms.
- Croatian sounds like poetry, English like prose. Meaning in prose is always clearer, more straightforward than in poetry.
- It is time to admit it: We have lost the battle against the English language, but we are winning the war for (global) communication.

### COMMENTS BY PEOPLE WHO CHOSE THE CROATIAN/CROATICIZED BLOG LANGUAGE:

- Whenever there is a Croatian translation for an English term, it should be used (response by a subject who studies Croatian as her A major).
- National languages should be defended against the invasion of English on the Web (comment by a respondent who proposes 'John Doe' as his blog nickname).
- I prefer the Croatian version of Web language because my English is not good, I do not feel comfortable with English, I frequently do not understand content expressed in English and also cannot express everything I want to in English.

## **5. Discussion and Conclusion**

A series of observations can now be made.

- Studies 1 and 2 both suggest that anglicized blogs have a better reception by the (blog) public than croaticized sites (higher number of hits; higher percentage in Questions 1, 2, and 8, as well as the choice of blog nicknames clearly tending towards English and anglicized options). This is particularly true for the categories 'up to date' (Question 4) and 'clarity' (Question 7), whereas it is less true for the category 'site usefulness' (Question 6). Curiously, a category that appears close to 'site usefulness' or 'site interest' is clearly biased towards the anglicized site (57.89% of respondents rate the anglicized site as being more interesting than the croaticized one).
- Analysis of the open-ended Question 9 reveals a number of tendencies: a) whenever there is a choice in usage between English and Croatian Internet-related words, most subjects prefer the former; b) subjects who do prefer the croaticized options motivate their choice exclusively in terms of language barriers (inadequate command of the English language). The one subject whose choice of the croaticized version was based on (Croatian) cultural identity preservation curiously opts for an English blog nickname, 'John Doe'. This example suggests that a great gap divides linguistic theory or policy and linguistic practice. Likewise, the open commentaries on the blog, while representing an unsystematic collection of individuals' thoughts, clearly suggest that the issue of croaticization vs. anglicization is not just a (Web) reality, but that, crucially, usage or practice, with its clear preference for the international/anglicized version, sharply contradicts the national linguistic recommendations. However, we need to strictly point out that these tendencies are those expressed *by* Internet users *about* Internet/blog language. As already pointed out in the

introduction to this paper, this type of language (just like the language of the fashion or music industry) represents a particular linguistic subcode, a jargon, and is thus likely to a) reflect only very specific tendencies by specific language users relative to a specific field, and b) require separate consideration within language policy mandates, with separate (possibly different) solutions than for the standard.

- In terms of discourse, blogs present a very interesting case of a total breakdown of the public-private barrier. This has yielded a new type of discourse strategy in terms of lexis, syntax and text exposition, coherence and cohesion structures, genre and code switching, among others. In this context we also note the peculiar style and discourse shifts when authored blogs (signed with the full name allowing identification of the author) are contrasted with blogs posted under a nickname (identity of the author unknown and untraceable).
- Finally, we wonder whether blogspeak (in our case, the blogspeak of university students, who will probably at least in part constitute tomorrow's societal elite) will influence the wider social discourse, and if so, how it will do so. Furthermore, given that the Internet and Internet language is outside the control of anyone, and given that linguistic state policy mandates do not apply in this case, we also wonder whether the peculiarities of the phenomenon are of the magnitude to merit research within a separate, new subfield: Internet Linguistics. Within this context the status of Internet jargon, as well as its relations to the standard, would need to be analysed and explained.

In view of all the above, we must conclude that economic, technological, political, social and other developments have brought about a world that is increasingly not just multilingual, but also cross-lingual. Field by field (rather than geographic area by geographic area), languages in contact create new varieties that then spread. In the case of the Internet, the universal language is without any doubt English (cf. Maurais, 2003), or, perhaps more correctly, it is a clumsy, misspelled English, a sort of 'lingua franca', international(ized) English.

We do agree with Thomas (2000) when he notes that not all national languages will necessarily be marginalized by English on the Web (cf. also the results of the study by Vehovar *et al.* (1999) of Slovenian vs. English on the Web). However, we maintain that the spread of English, and in particular its 'internationalization', might indeed be paralleled by another process of 'internationalization' that manifests as, or rather translates into, the 'anglicization' of national languages. In other words, the internationalization of English goes hand in hand with the anglicization of other languages. Furthermore, the two processes could be mutually defining, for both contain elements of the language-identity issue and bear on language policy.

What the implications of these phenomena are for national cultures and national identities, and for language policymakers, is the focal question of the final part of this paper.

Albeit diminishing, the hegemony of English in the virtual space is more than likely to continue to be felt for a long time (see Maurais, 2003). This also means that the use of English does and will continue to polarise the world into Internet users and Internet illiterates, at least to an extent<sup>11</sup> (see the second point in the above discussion). Curiously, the same logic can be applied to other national languages: The degree of openness to the non-standard national Web language – which also represents the 'standard of the Web', some type of net-jargon or 'netspeak' – equates to being Internet literate in that language.

While partially penalising for some, this perception of Internet literacy does have a positive corollary effect, for it can be viewed as a motive for language expansion. It is true that the trend of global language expansion does in practical terms more often than not translate into the expansion of English (and, to an extent, into its 'distortion'), namely into the anglicization of other (minor) languages (i.e., 'distortion' with respect to their standard). Moreover, some view the global trends described in this paper as bringing about the impoverishment of (smaller) national languages and cultures in general. Here, however, we suggest a more optimistic view: That the phenomena of languages and cultures in contact, or rather, the features resulting from that contact, can also be seen as a development of *both* the guest (usually English) and the host national language. In such a case, we could conclude with Crystal (2006) that

the Internet is going to record linguistic diversity more fully and accurately than was ever possible before. What is truly remarkable is that so many people have learned so quickly to adapt their language to meet the demands of the new situations, and to exploit the potential of the new medium so creatively to form new areas of expression. (...) The arrival of Netspeak is showing us *homo loquens* at its best. (ibid: 276)

We may conclude at this point that in view of the above, language policy mandates should become more flexible with regard to (blindly) defending natural cultures or (linguistic) identities, particularly where specific jargons are concerned. Today, perhaps more than ever before in human history, forecasting but also controlling the fate of languages or linguistic evolution is most uncertain, if not impossible. It is indeed true that challenges posed by the continuing spread of English affect all fields, but it is also true that a reliable model for the prediction and control of the expansion process needs to reflect the multidimensional and multifunctional nature of language dynamics (e.g.,

---

<sup>11</sup> Suffice it to say that the percentage of Web pages by language in 2000 was 68.4 English, followed by 5.9 Japanese and 5.8 German (see Maurais, 2003).

economic, psychosocial, cultural, historical and other factors – cf. Mackey, 2003: 70-76), which is not always possible. We might best accept that a certain degree of change is unavoidable, unstoppable, and not necessarily bad (or, at least, not subject to evaluation), but just a record of the times at hand, and that change represents a reality perhaps better dealt with from a descriptive, rather than from the prescriptive, perspective.

In the case of Croatian, we maintain that a successful language policy needs to acknowledge and accept the fact that narrow promotion of political (which sometimes overshadow the national) interests by means of language is a negative practice that lacks both benefits for and reception by language users. The data reported in this study paint one picture of reality that should be taken into consideration by language policymakers in Croatia (and possibly elsewhere), where currently this seems not to be the case. Language policy cannot blindly insist on dismissing English and anglicized terms and on promoting national linguistic choices that are not used by the general public, let alone by the Internet users, when reality points to English and anglicization as two unstoppable, global trends. In other words, language policymakers, who currently strongly insist on a top-down approach to language policy, should consider a bottom-up approach. In this context, the top-down approach can be equated with the strictly prescriptive tradition, where language behaviours are imposed onto users, leaving no room for flexible approaches to at least some jargons. Conversely the bottom-up approach can be equated with the descriptive tradition, which leaves much room for usage practices that constantly interact with the norm, shaping it to a degree.

For as much as language is a code or a system of rules, it is also (if not foremost) a behaviour, a world of developing practices. While we may describe and prescribe codes, it is far more difficult to confine behaviours, since behaviours, including linguistic ones, tend to trespass the borders meant to confine them, ultimately bringing about new codes. This reality cannot be ignored by anyone wishing to put linguistics at the *service* of language users.

Two conclusions can be reached from the above analyses and discussions. First, when talking about the language-identity-policy trinomial, we need to distinguish between two linguistic realities, namely (Internet) jargon vs. standard. Next, with regard to these linguistic realities, finding the fine line between the individual and the social on one hand, and between singular national interests and global developments (understanding and allowing for 'global evolution') on the other, is now and for many years to come among the most difficult and most important tasks for linguists. This also means that today's language policymakers must constantly search for a perfect balance between the national and the global. One goal of this paper has been to propose that this balance is unreachable unless we also consider the individual, psychosocial level of language usage. A language policy is useless

unless it can be successfully implemented. Paramount to this implementation process is understanding how and why people react to language policies, which means understanding the linguistic perceptions and preference mechanisms of individuals such as those of the Croatian blog users illustrated in this paper.

## BIBLIOGRAPHY

- Brala, M. (2007): I am the language I speak: Bilingualism and identity. In: HDPL 2006 Proceedings. Split.
- Crystal, D. (2006): Language and the Internet. Cambridge (Cambridge University Press).
- Dewale, J.-M. (2005): Investigating the Psychological and Emotional Dimensions in Instructed Language Learning: Obstacles and Possibilities. In: The Modern Language Journal, 89, 367-380.
- Dewaele, J.-M. & Pavlenko, A. (2001/2003): Webquestionnaire Bilingualism and Emotions. London (University of London).
- Greenfield, P. & Subrahmanyam, K. (2003): Online Discourse in a Teen Chatroom: New Codes and New Modes of Coherence in a Visual Medium. Journal of Applied Developmental Psychology, 24, 713-738.
- Grosjean, F. (1982): Life with Two Languages: An Introduction to Bilingualism. Cambridge MA (Harvard University Press).
- Herring, S. C. *et al.* (2004): Bridging the gap: A genre analysis of weblogs. Proceedings of the 37<sup>th</sup> Hawaii International Conference on System Sciences (HICSS-37). Los Alamitos, CA: IEEE. Retrieved April 4<sup>th</sup> 2006 from <http://csdl.computer.org/comp/proceedings/hicss/2004/2056/04-205640101b.pdf>
- Huffaker, D. (2004b): Spinning yarns around a digital fire: Storytelling and dialogue among youth on the Internet. First Monday, 9 (1). Retrieved April 4<sup>th</sup> 2006 from [http://www.firstmonday.dk/issues/issue9\\_1/huffaker/index.html](http://www.firstmonday.dk/issues/issue9_1/huffaker/index.html).
- Huffaker, D. A. & Calvert, S. L. (2005): Gender, Identity and language use in teenage blogs. Journal of Computer-Mediated Communication, 10 (2).
- Japundzic, A. (2006): Stanje blogosfere. Rezultati ankete (The state of the blogosphere. Questionnaire results). Retrieved June 3<sup>rd</sup> 2007 from <http://marketingservis.com/blogosfera/stanje-blogosfere-anketa/>.
- Mackey, W. (2003): Forecasting the fate of languages. In: Maurais, J. & Morris, M. A. (2003): Languages in a Globalising World. Cambridge (Cambridge University Press).
- Maurais, J. & Morris, M. A. (2003): Languages in a Globalising World. Cambridge (Cambridge University Press).
- Maurais, J. (2003): Towards a new global linguistic order? In: Maurais, J. & Morris, M. A. (2003): Languages in a Globalising World. Cambridge (Cambridge University Press).
- Pavlenko, A. (2006): Bilingual selves. In: Pavlenko, A. (ed.): Bilingual minds. Emotional experience, expression, and representation, 1-33. Clevedon (Multilingual Matters).
- Starcevic, A. & Geld, R. (2007): Hrvatski standardni jezik i engleski utjecaj (The Croatian standard language and the English influence). Paper presented at the HDPL 2007 conference. Split, Croatia.
- Thomas, N. (2000): How much IT can minority languages afford? Editorial, Contact 16 (3), 2.
- Vehovar, V., Batagelj, Z. & Lozar, K. (1999): Language as barrier. Internet Society Proceedings. [www.isoc.org/inet99/proceedings/3i/3i\\_3.htm](http://www.isoc.org/inet99/proceedings/3i/3i_3.htm).

- Warschauer, M. (2000): Language, Identity, and the Internet. In: Kolko, B., Nakamura, L. & Rodman, G. (eds.): *Race in Cyberspace*. New York (Routledge). 151-170.
- Wright, S. (2004) MOST and Initiative B@bel funded research: Language on the Internet. UNESCO's *International Journal on Multicultural Societies*. Vol. 6, N° 1.
- Wilson, R. (2005): Another language is another soul. Paper presented at the EUROSLA 2005 conference. Dubrovnik, Croatia.