Michał Głuszkowski
Nicolaus Copernicus University

Creolization and Balkanization as a Result of Language (Dialect) Contact. Is the Origin of Mixed Languages universal?

This article is an attempt to find some common points in the development of languages in various contact situations and theories used to describe them. Scholars investigating the situation of Balkan and creole languages have developed tools to analyze the phenomena of interference and convergence, which can also be used in other fields of contact linguistics. Thus, even linguists who are interested in other types of the language contact cannot be indifferent towards the theoretical achievements of these disciplines. Independently of the subject of research, scholars investigating languages in contact cannot avoid defining certain basic phenomena and Uriel Weinreich’s theoretical assumptions. Basing ourselves on his main definitions of language contact⁴,

⁴ “Two or more languages will be said to be in contact if they are used alternately by the same persons” (Weinreich, 1963, p. 1).

This work was supported by the Polish Ministry of Science and Higher Education.
Competing interests: no competing interests have been declared.
Publisher: Institute of Slavic Studies PAS.
This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 PL License (creativecommons.org/licenses/by/3.0/pl/), which permits redistribution, commercial and non-commercial, provided that the article is properly cited. © The Author(s) 2015.
bilingualism\textsuperscript{2} and interference\textsuperscript{3} we may classify language contact situations according to several different criteria. Not less important than these terms is U. Weinreich’s approach to the individual and group dimension of language contact. The relationships between two or more communities are usually initiated only by some of their members. Interference and language changes occur at first in their own speech and only after that are they adopted (or not) by the rest of the community (cf. Głuszkowski 2013, pp. 38–40; Weinreich, 1963, p. 83). One has to consider this fact while characterizing the type of language contact.

According to the main features of Balkan and creole languages, as well as other language situations of interest to us in this article, we will focus on one criterion of classification – the symmetry of influence. Thus, symmetrical and asymmetrical relationships between languages in contact can be distinguished. In symmetrical configurations both languages involved in the contact situation exert relatively equal influence on each other and we cannot indicate which one is the source and which is the target language. This situation may be observed in the contact of Balkan languages. Jouko Lindstedt proposes that:

“Linguistic balkanization was initiated by speakers who were bilingual or multilingual to such an extent that in their speech there were transfers not only from, but also into their native languages and who for that reason favoured features that made it easier to identify structures across language” (Lindstedt, 2000, p. 241).

Following Maria Todorova he also notices that the term of “balkanization” in linguistics refers to convergent development and is understood differently than in common language, where it is associated with division and disintegration (Lindstedt, 2000, p. 234).

On the contrary, in asymmetrical contact roles of the languages involved are determined and we can distinguish the source language as well as the target one\textsuperscript{4}. In such a relationship the source, i.e. the dominant and usually majority language has an influence on the target language which is usually a minority

\textsuperscript{2} “The practice of alternately using two languages will be called bilingualism, and the persons involved, bilingual” (Weinreich, 1963, p. 1).

\textsuperscript{3} “Those instances of deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language, i.e. as a result of language contact will be referred to as interference phenomena” (Weinreich, 1963, p. 1).

\textsuperscript{4} One has to notice that the roles are not determined absolutely and, depending on the socio-cultural conditions of the contact, the target language may also affect the source, i.e. roles can be swapped.
language or has lower social status. In these configurations the community using the target language is predominantly bilingual and the users of the source language – monolingual. According to William Mackey, such a situation may result in language shift. A fully bilingual group “in which everyone is fluent with two languages has no reason in remaining bilingual while they can get along with one of their languages” (Mackey, 1968, pp. 554–555).

The situations of asymmetrical language contact may result in the processes of pidginization and creolization. John Holm defines pidgin as:

“[… ] a reduced language that results from extended contact between groups of people with no language in common; it evolves when they need some means of verbal communication, perhaps for trade, but no group learns the native language of any other group for social reasons that may include lack of trust or close contact. Usually those with less power (speakers of substrate languages) are more accommodating and use words from the language of those with more power (the superstrate), although the meaning, form and use of these words may be influenced by the substrate languages” (Holm, 1988, pp. 4–5).

J. Holm claims that pidgin is a more stable code than jargon and that this stabilization requires hybridization and that pidgin may be adopted by the users of substrate language for internal communication, which supports W. Mackey’s assumptions (cited above) and coincides with David Crystal’s observations, who used the notions of “makeshift”, “marginal”, and “mixed languages” to characterize pidgins (cf. Al-Jasser 2012, p. 71; Holm, 1988, p. 5; Mackey, 1968, pp. 554–555).

Despite the differences between world pidgins which emerged from the contact of a colonists’ language (English, French, Portuguese, Spanish, German, Dutch and others) and a language of the locals or any community with lower social status, they have many common structural features. Mark Sebba made an attempt to summarize them. Some of his conclusions directly refer to the subject of this article: pidgins are usually based predominantly on the vocabulary of one of the languages in contact, their grammars are reduced and more simplified than the grammars of the input languages, their vocabularies are limited and words cover a wide semantic range, and their morphology tends to analytism and agglutination (Al-Jasser 2012, pp. 72–73).

The next stage of pidginization is creolization. Creole language

“[…] has a jargon or a pidgin in its ancestry; it is spoken natively by an entire speech community, often one whose ancestors were displaced geographically so that their ties with their original language and sociocultural identity were partly broken” (Holm, 1988, p. 6).
As has been mentioned in most studies on creole languages, their origin is connected with colonization and slavery (Hesseling, 2009; Holm, 1988, p. 6). The community using a creole language do not have to be bilingual – the creole language may be their only means of communication.

The processes of pidginization and creolization result in the emergence of mixed languages (codes). However, this is not their only source. In some situations the total amount of interference in a language of lower social status in a bilingual society reaches the level at which lexical, syntactical, morphological and phonetic features from both codes are mixed up and the minority language becomes a hybrid. The hybridized system is used in parallel with the dominant language. A special example of a mixed code on the basis of the language of lower prestige is Belarusian trasyanka. On the basis of Belarusian dialects under the influence of the Russian language, a new mixed code has developed – with predominantly Belarusian grammar and predominantly Russian lexis (see e.g. Hentschel, 2008). Similar processes may be observed in the contact of Ukrainian and Russian, which resulted in the creation of another mixed code – Surzhik (Del Gaudio, 2006, pp. 235–236).

There are also other paths which lead to mixed codes. Some bilingual communities tend to switch and mix the codes used by them. As a result of mixing of codes a fused lect may emerge (Auer, 1999). The new mixed code is a form of expression of the group identity or at least of a part of the group as the “we-code” (Jørgensen, 1998, p. 242). It has been observed, for example in the young generation of Italian and Turkish immigrants in Germany described by Katja Cantone and Volker Hinnenkamp (Cantone, 2007; Hinnenkamp, 2003). The new code built on the basis of two languages, e.g. Italian and German may be used in the minority community parallel either with both initial languages, i.e. German and Italian (the community is then in fact trilingual) or only with one of them – the dominant language of the society.

While the situations described above are a result of asymmetrical language contact, we can either point out the source and the target language or determine which language has been replaced by the new code. The language with higher social status remains practically intact.

Thus some differences between the processes of balkanization on one hand, and pidginization and creolization on the other, can be seen at first glance. As a result of symmetrical (or relatively symmetrical) language contact, balkanization first of all bore fruit with the presence of common features in different languages in one area. Of course, when we calculate the indices
of Balkanization of different languages and language groups involved in the process of Balkanization, we will get different scores dependent on the balkanisms observed in different languages – Slavonic languages are more balkanized than other groups and the Romani language is less balkanized (Lindstedt, 2000, p. 234). However, the influence is mutual and multidirectional. Not one of the languages from the Balkan Sprachbund remains intact.

What justifies the question about mixed languages and creolization in the Balkans then? The advocate of cooperation between Balkan linguistics and creolistics – because of the similarities of the subject of both subdisciplines – was Uwe Hinrichs. He enumerated several concurrences between Balkan and creole languages. According to his most significant observations, creole and Balkan languages represent the analytic (isolating) type of languages and both disciplines refer to not directly cognate languages which change their forms in the process of language contact (Hinrichs, 2002, pp. 149–151, 2004, pp. 147–162). However one has to underline the differences:

a) Neither pidgins nor creoles have developed in the Balkans.

Peter Bakker, one of the leading scholars in the field of mixed languages, claims that “the term ‘mixed language’ has been incorrectly applied in the past to all kinds of languages with some visible influence from other languages” and demands that we “limit its use to those cases where genetic classification is no longer possible” (Bakker, 2000, p. 29). In spite of numerous common features – balkanisms, the Balkan languages are still perceived as different codes. We are able to recognize if someone is speaking Macedonian or Albanian, while it is impossible to determine whether Haitian creole is more French or Western African. The level and range of convergence is simply not comparable either to creole languages or to mixed codes such as trasyanka or surzhik.

b) The source and the target language cannot be determined in Balkans.

J. Lindstedt claims that it is impossible to find a single source language for many grammatical balkanisms:

“The source language simply does not exist in the traditional sense: the sociolinguistic contact situation has caused changes that would not have occurred in any of the Balkan languages by internal drift” (Lindstedt, 2000, p. 238).

As J. Lindstedt pointed out, most of the grammatical Balkanisms would not appear in the internal development of any of the languages in the Balkans, whereas they are the result of a multilingual contact situation, and the notions
of “source language” and “target language” in their traditional meaning should not be used there (Lindstedt, 2000, p. 231). In creole languages the source and target are socially determined from the beginning.

c) The configuration of languages-in-contact prestige is different in the situations in question.

Although some creoles, e.g. Jamaican creole and Haitian creole have gained in prestige, it is still incomparable to the social position of English or French. In the Balkans the situation is different. Of course, there are some differences. J. Lindstedt describes various positions of Balkan languages on the scale of prestige. According to his socio-historical analysis, Greek has always been close to the top, while Romani has been near the bottom (Lindstedt, 2000, pp. 239–242). Both codes are less balkanized than Macedonian, Bulgarian, Albanian, and Aromanian, which are near the middle of the prestige scale, but one has to stress that balkanisms appear in all the languages of the Balkan Sprachbund, independently of their prestige (Lindstedt, 2000, pp. 240–243). In other (not Balkan) situations of language contact, languages of higher prestige are usually immune to the outer influence while those of low prestige – are susceptible to influence.

d) Social consciousness and the bilingual community attitude towards their language(s) are different.

When a mixed language replaces the minority language or the input languages, the speakers are often not aware of the language change taking place. A lot of people in Ukraine and Belarussia speak trasyanka or surzhik without realising it. Neither they nor observers from outside are able to distinguish elements of two mixed codes without a comprehensive analysis. The situations in the Balkans is different and J. Lindstedt described it as stable multilingualism and after John Gumperz and Robert Wilson enumerated the most important conditions of this state: speakers of different languages live closely together, there is no lingua franca, all groups have sufficient access to other languages, and native languages are important symbols of group identity (Lindstedt, 2000, p. 239). J. Lindstedt underlines, that in such a situation “people are still perceived as speaking different languages”, which amounts to an important difference between Balkan languages and mixed codes such as creoles or trasyanka and surzhik (cf. Lindstedt, 2000, p. 239).

Balkan languages, creoles and pidgins, languages in multilingual societies as well as language islands enjoy various different types of language contact,
which results in the differences between them. However, even in very similar language situations, in the frames of the same type of language contact, bilingualism may develop in different ways. Thus I would like to compare bilingualism of two migrant communities living in foreign surroundings for many generations: Russian Old Believers in Poland and Poles in Siberia.

The Old Belief movement began in the second half of the 17th century after the reforms in the Russian Orthodox Church. As a great number of Orthodox people had not accepted the reforms and were deprived of church and civil rights, they had to migrate. A lot of them escaped to the Commonwealth of Poland and Lithuania. Nowadays the Old Believers in Poland live in a language island. Most of the community members are bilinguals but until the beginning of the 20th century, due to the isolation of their communities, they were predominantly monolingual (especially women).

They acquired the language of the dominant group (Polish), but it was still important for them to preserve the language of their ancestors (a Russian dialect from the Pskov-Novgorod group). Their bilingualism is connected with diglossia, which also affects language choice and the phenomenon of code switching and code mixing. Due to the amount of interference Stefan Grzybowski and Dorota Paśko-Konczniak described the present state of Old Believers dialect as a hybridized system. The hybrid is based on Russian morphology, which form the basis (morphological frame) for other elements coming from both codes. The new mixed system is replacing the dialect in its traditional form.

The second bilingual community to be compared is Vershina, a Polish village in Siberia, 130 km northwards from Irkutsk. It was founded in the beginning of the 20th century by Polish voluntary settlers from Little Poland.

---

5 According to Peter Rosenberg: “internally structured settlements of a linguistic minority on a limited geographical area in the midst of a linguistically different majority” (Rosenberg, 2005, p. 221).

6 One has to recognize that the Old Believers community still uses the Church Slavonic language in liturgy. It is not used for communication, but is involved in their diglossia (poliglossia).
In the first years after migration the inhabitants of the village preserved their Polish dialect from Little-Poland, traditions, farming methods and machines and also the Roman Catholic religion. The years of sovietization which in fact started together with collectivisation in the 1930s entailed serious social changes in the community, which also affected the language situation of the group. Nowadays Vershina constitutes a bilingual community; the ancestors of Polish settlers still use a mixture of local dialects from their place of origin and the Russian language. In spite of the interference observed in the Polish dialect, the system remained relatively clean – we can speak neither of a hybrid nor of a mixed code. The inhabitants of Vershina mainly preserved the Polish dialectal phonetic features and the loan translations, i.e. influence in morphology and syntax are limited (cf. Głuszkowski, 2011b; Paśko, 2009). Practically the only visible interference refers to the field of lexis. However, due to the sociocultural conditions, the Polish dialect is vanishing. This situation may be described as the occurrence of language shift rather than convergence.

Table 1. Analogies and differences between two Slavonic language islands.

<table>
<thead>
<tr>
<th>Old Believers in Poland</th>
<th>Poles in the village of Vershina (Siberia)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>analogies</strong></td>
<td></td>
</tr>
<tr>
<td>language island</td>
<td>language island</td>
</tr>
<tr>
<td>mother tongue: Russian dialect, i.e. non-standardized language variety, without written form</td>
<td>mother tongue: Polish dialect, i.e. non-standardized language variety, without written form</td>
</tr>
<tr>
<td>Eastern-Western Slavonic language contact</td>
<td>Western-Eastern Slavonic language contact</td>
</tr>
<tr>
<td>contact with a national, standardized language with a literary and cultural heritage (Polish)</td>
<td>contact with a national, standardized language with a literary and cultural heritage (Russian)</td>
</tr>
<tr>
<td>Polish language plays the role of literary variety for the Old Believers dialect (<em>hochsprachliches Dach</em>)</td>
<td>Russian language plays the role of literary variety for the Polish dialect (<em>hochsprachliches Dach</em>)</td>
</tr>
<tr>
<td>religious minority: Old Believers among Roman Catholics</td>
<td>religious minority: Roman-Catholics among Orthodox Christians, atheists and shamanists</td>
</tr>
<tr>
<td>rare cases of interference in their Polish</td>
<td>very rare cases of interference in their Russian</td>
</tr>
</tbody>
</table>

* Heinz Kloss’ notion of *hochsprachliches Dach*, i.e. the literary variety of the language L₁ for the non-literary variety of the language L₂, used after Gerd Hentschel (see: Hentschel, 2002, p. 85).
<table>
<thead>
<tr>
<th><strong>Old Believers in Poland</strong></th>
<th><strong>Poles in the village of Vershina (Siberia)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>differences</strong></td>
<td></td>
</tr>
<tr>
<td>language hybrid (a mixed code)</td>
<td>bilingualism; interference in mother tongue; language shift instead of code-mixing</td>
</tr>
<tr>
<td>biculturalism</td>
<td>cultural hybrid</td>
</tr>
<tr>
<td><strong>possible reasons for different development</strong></td>
<td></td>
</tr>
<tr>
<td>strong tendency to isolation (weakened only in the last quarter of the 20th century)</td>
<td>lack of special tendency to isolation</td>
</tr>
<tr>
<td>relatively low tolerance to mixed marriages</td>
<td>high tolerance to mixed marriages with Slavonic partners (Russian and Ukrainian)</td>
</tr>
</tbody>
</table>

In the two language situations, the factors which seem to be analogical are the following: both communities form language and cultural islands, their mother tongue is a non-standardized language variety, without written form, they represent the Eastern-Western Slavonic type of language contact, and the language of their dominating surroundings is a standardized code with a literary and cultural heritage, the dominant language is their *hochsprachliches Dach*, both communities are religious minorities and experienced severe socio-cultural as well as political changes. However, similar sociocultural and language conditions led to different development of the island languages. According to the results of our investigations so far the differences are caused mainly by stronger isolation of the Old Believers community. These two examples show that not it is only in situations of language contact of different types – symmetrical, asymmetrical, multidirectional, in one direction etc. – that the genesis of mixed codes is not universal. U. Hinrichs’ statement that the subdisciplines of contact linguistics may benefit from each other cannot be undermined, but one has to underline that all cases of language and culture contact are in some degree unique and need an individual approach.

**BIBLIOGRAPHY**


**BIBLIOGRAPHY**

**(TRANSLITERATION)**


Michał Głuszkowski  Creolization and Balkanization as a Result of Language (Dialect) Contact…


Kreolizacja i bałkanizacja jako rezultat kontaktu języków (dialektów). Czy istnieje uniwersalna geneza języków mieszanych?

Możemy wyróżnić kilka typów kontaktu językowego w zależności od relacji pomiędzy językami. W artykule omówiono rezultaty kontaktu językowego z jedno- i wielokierunkowym wpływem: bałkanizacja, pidginizacja, kreolizacja oraz inne typy kontaktu, a także podjęto próbę odpowiedzi na pytanie o to, czemu tylko niektóre z nich skutkują powstaniem języków mieszanych. Analizie poddano różne podejścia teoretyczne, służące opisowi języków w kontakcie, procesowi konwergencji oraz powstaniu języków mieszanych. Ponieważ porównanie takich przykładów kontaktu językowego jak Bałkańska Liga Językowa, społeczności kolonialne i postkolonialne oraz wieloetniczne społeczeństwa na Zachodzie pokazało, że każdy typ kontaktu wymaga innego podejścia, końcową część artykułu poświęcono dwóm praktycznie analogicznym sytuacjom językowym: Rosjanom-starobrzędowcom w Polsce i Polakom na Syberii. Pomimo wielu podobieństw okazuje się jednak, że w obydwu społecznościach bilin­gwizm rozwinął się w inny sposób.

Słowa klucowe: kontakt językowy; kreoloizacja; bałkanizacja; kody mieszane; dwujęzyczność i dwukulturowość

Creolization and Balkanization as a Result of Language (Dialect) Contact. Is the Origin of Mixed Languages universal?

There are several types of language contact depending on the relations between languages. The article focuses on the results of language contact with multi- and unidirectional influence: balkanization, pidginization, creolization and other types of contact – why not all of them result in mixing codes. The author considers various theoretical approaches to describe languages in contact, the process of convergence and the genesis of mixed codes. While the comparison of such language situations as Balkan Sprachbund, colonial and postcolonial societies, multi-ethnic societies in the Western world, has shown that each type of language contact needs its own approach, the final part of this paper is devoted to two analogical language situations: Russian Old Believers in Poland and Poles in Siberia. However, despite of many similarities, even these two communities have developed their bilingualism in a different way.

Keywords: language contact; creolization; balkanization; mixed codes; bilingualism and biculturism
Notka o autorze

Michał Głuszkowski (micglu@umk.pl) – dr hab. nauk humanistycznych w zakresie językoznawstwa, adiunkt w Katedrze Filologii Słowiańskiej Uniwersytetu im. Mikołaja Kopernika w Toruniu. Zainteresowania badawcze: socjologiczne i psychologiczne aspekty dwujęzyczności, metodologia badań lingwistycznych, funkcjonowanie gwar wyspowych w warunkach bilingwizmu, dialektologia rosyjska i polska, historia socjologii polskiej i rosyjskiej.

Michał Głuszkowski, PhD (micglu@umk.pl) – linguist, assistant professor at the Nicolaus Copernicus University in Torun, head of the Department of Slavic Studies. Research interests: sociological and psychological aspects of bilingualism, linguistic research methodology, language and cultural islands in the conditions of bilingualism, Russian and Polish dialectology, history of sociology, Polish and Russian.