

Mutual Intelligibility among the Slavic Languages

by Robert Lindsay

The mutual intelligibility (MI) of the languages of the Slavic family is an interesting topic because many are mutually intelligible to one degree or another. It is commonly believed that all Slavic languages are fully mutually intelligible, which implies that they are close enough that they can fully communicate with each other easily. This is simply not the case, though there is a grain of truth to the myth as with most stereotypes. This is because unlike say English and the rest of Germanic, there is usually "some" intelligibility between one Slavic language and another.

Overview

It is important to note that the idea of this paper was try to test "pure inherent intelligibility." A pure inherent intelligibility test would involve a couple of things.

First it would involve a speaker of Slavic Lect A listening to a tape or watching a video of a speaker of Slavic Lect B so as to control for the extralinguistic additive intelligibility that comes with face to face communication. This would account for the "pure" part of the test.

In order to account for the "inherent" part of the test, it would also be best to find "virgin ears" as much as possible when testing MI, as bilingual learning can be a serious problem that throws off intelligibility figures in the direction of overestimates. Intelligibility without any prior bilingual learning is regarded as "inherent" intelligibility. Bilingual learning is a significant confounding variable with testing intelligibility in a number of Slavic languages, particularly closely related languages like Czech and Slovak that have a lot of interaction.

Written intelligibility is often very different from oral intelligibility in that in a number of cases, it tends to be higher, often much higher, than oral intelligibility. Written intelligibility was only calculated for a limited number of language pairs. Most pairs have no figure for written intelligibility.

Method

First, a Net search was done of forums where speakers of Slavic languages were discussing how much of other Slavic languages and lects they understand. An attempt was made to exclude speakers with a lot of exposure to the other lect in order to weed out bilingual learning. These reports were then tallied and averaged together.

Second, a number of native speakers of various Slavic lects were interviewed about MI with other Slavic lects, language/dialect confusion, the state of their language, its history and so on.

Third, I played native speakers of Slavic Lect A videos of Slavic Lect B and had them estimate how much they understood. Similarly, I showed speakers of Slavic Lect A writing samples of Slavic Lect B and had them estimate how much they understood. Separate figures were obtained for oral and written intelligibility.

Fourth, linguists and other Slavic language experts were interviewed to obtain MI figures for various Slavic lects.

Finally, I researched the small number of formal intelligibility studies on Slavic. However, formal Slavic intelligibility studies have been very few in number, and the area is only beginning to be seriously researched in the past few years.

Some reports were simply discarded. For instance, a few reports came in from ethnic nationalists and ultranationalists with an obvious nationalist agenda. Ethnic nationalists typically say that there is only one language in their country, and that is the national language. They claim that all other lects spoken in the country are dialects of the national language. They tend to vastly overestimate MI with languages related to their own.

Some who have agendas outside of their nation and wish dominion

over large areas of land and people also claim that they can understand many related tongues of the same family. For instance Turkish nationalists typically claim that Turks can easily understand all other Turkic languages and that all Turkic languages are all dialects of Turkish. They say this because they think that all speakers of Turkic languages are part of a fantasized vast Turkic nation that would ideally cover the entire area where Turkic lects are spoken.

All reports coming from a nationalist, ethnic nationalist, or pan-nationalist perspective were discarded due to possible bias.

The percentages quoted by Net searches of statements by native speakers, native speaker interviewees, and linguists and language teachers in Slavic countries lined up fairly well.

Figures were weighted in four ways:

Fourth Level weighting was given to intelligibility figures which were tallied for all language pairs from native speaker sources. The best were where I had multiple figures which could be averaged together. These figures were tallied up for each pair of languages and were then averaged together. In the worst cases, I only had a figure from only one informant. Hence all figures in this group are averages of judgments taken from statements by native speakers of the languages in question.

Tertiary weighting was given to a few cases where I was able to give a speaker of Lect A a recorded or written sample of Lect B to listen to it or read. I then had the informant estimate how much they understood of the sample. A result here would often outweigh the fourth-level judgments.

Secondary weighting was given in cases where I had a figure from a Slavic linguist, language professor, university graduate, or grad student who either had a Linguistics-related degree, was working on one, or who had extensive background in Linguistics courses or study. The figures from experts were derived via both academic works and personal communications. These expert judgments outweighed the previous two judgments.

Primary weighting was given to the three cases of intelligibility testing, a one-way test with four Slavic languages (Sloboda and Brankaikec 2012), and a five-way test of six Slavic languages (Golubovic and Gooskens 2015a) and a one-way test with three

languages (Majewicz 1986). These results outweighed the previous three determinations.

One might wonder how linguists normally determine intelligibility among pairs of lects. This is simple.

Linguists usually simply ask native speakers of Lect A how much they understand of Lect B. All of these native speaker estimates are then averaged together.

They may also ask linguists who speak or study Lect A or B for an MI figure.

In some cases, formal intelligibility testing is done (Casad 1974). However, the vast majority of the languages in the world, including many large languages, have not undergone intelligibility testing. But this is beginning to change in recent years, as intelligibility testing has become something of a "hot" topic, especially in Europe. Obviously, formal intelligibility testing is the best method of all for determining MI.

Most MI figures obtained from a linguist are probably estimates derived from native speaker knowledge and not a result of formal testing.

Testing by the Summer Institute of Linguistics has shown that intelligibility testing has high validity and reliability (Casad 1974). These results refute the common criticism that MI is not measurable in any way using any technique.

Historically, formal testing has not often done *with major languages*. In addition, most large languages have historically not been tested often or even at all for MI.

However, this seems to be changing due to a European multi-university project out of the University of Groningen, Netherlands called MICRela, Mutual Intelligibility of Closely Related Languages (Golubovic and Gooskens 2015a, Golubovic and Gooskens 2015b, Golubovic and Gooskens 2014a, Golubovic and Gooskens 2014b, Golubovic and Gooskens 2014c, Golubovic 2013, Swarte et al 2012).

Funding has been provided for this project from a number of sources, and studies are now going on in the Germanic, Slavic, and Romance language families.

Slavic MI studies have already been done by this group (Golubovic

and Gooskens 2015a, Golubovic and Gooskens 2015b, Golubovic and Gooskens 2014a, Golubovic and Gooskens 2014b, Golubovic and Gooskens 2014c, Golubovic 2013, Swarte et al 2012).

A problem arises in that even formal intelligibility testing has a wide scatterplot ratio. For instance, two short studies of Turkish intelligibility of Azeri found an intelligibility range of between 66-92% (Roos et al 2009). A third test of Turkish-Azeri intelligibility found a much lower level of understanding at 49% (Fraenkel 1962). So three studies of Turkish-Azeri intelligibility found figures ranging from 49-92%, with an average across all three of 69%.

It is true that there is some skew between MI estimates of native speakers and formal intelligibility testing.

There has been one study testing skew in Slavic MI. In this study, speakers' estimates varied an average of only 13% from testing results. The 13% skew was due to both overestimates and underestimates in this study (Sloboda and Brankaikec 2012).

In order to better estimate skew between native speaker estimates and formal intelligibility testing in Slavic, we should conduct more studies, as one study is not enough to form much of a conclusion on.

Limitations

The first limitation in this study is that it relies heavily on self-reports from native speakers on how much they understand of the other language. Although this would seem to be an excellent way to study this question, the problem is that speaker reports can be off by quite a bit.

Native speakers will vary in their ability to understand a closely related language based on a variety of factors. Reasons for the discrepancy are not known.

Some people simply are better able to understand a closely related language than other people.

People with higher intelligence and more education may be more able to pick out cognates and decipher differences between closely related languages better than people lower intelligence and less education.

People who understand more than one language in a family may be better able to understand other languages in the family.

Exposure is also important. Ideally, MI studies should look for "inherent intelligibility." This means L1 speakers who have been exposed to little to none of the L2. This shows what pure MI between the languages is without throwing in the conflating factor of "bilingual learning." Bilingual learning has the potential to seriously skew MI studies of inherent intelligibility.

Although MI studies testing for inherent intelligibility often try to find L1 speakers with "virgin ears" who have been exposed to little of the L2 they are testing, in practice this is difficult to do, especially in cases like Czech and Slovak, whose speakers have heavy exposure to each other for most of their lives. In these cases, we are often measuring MI as it exists on the ground in reality as opposed to inherent MI, which might be measurable more in theory than in reality here.

Because of bilingual learning, knowledge of more than one language in a family, intelligence, education, or language learning ability, there is often considerable skew in native speaker reports. Nevertheless, these reports can be quite accurate if you gather enough subject reports and average them together. With enough subjects averaged, factors such as differential abilities, bilingual learning, intelligence, and knowledge of related L2's will tend to smooth out as you get subjects who vary widely in these factors.

In other words, speaker reports do show some scatter and skew. Not uncommonly they are quite close to the findings of formal studies, but in other cases they can be considerably off.

Speaker reports of written MI tend to be off by the most. Speakers commonly vastly overestimate how much written MI they have of the other language. Formal MI studies often show that speakers understand much less of the written L2 than they think they do.

Linguist or expert judgments are a good source of information, as they are often given cautiously by linguists who have been studying the language for a while, typically years. Judgments given by linguists track better with formal MI results than speaker reports.

Nevertheless, comparisons between estimates by linguists and formal intelligibility studies show that even expert judgments can be off,

typically not by a lot (Golubovic and Gooskens 2015a) but sometimes by up to 50 percentage points (Sussex and Cubberley 2011)!

Formal MI studies should be regarded as the Holy Grail in MI research. Assuming they are done properly, the figure arrived at should be regarded by science as the last word on the subject for the time being. In other words, the statement should be, "Formal scientific intelligibility testing has shown that Language X has an average of 69% intelligibility of Language Y."

If others wish to test the findings to attempt to validate them, they may do so, but this is usually not done, as most linguists I have spoken to do not think MI research is worthy of controversy. In cases where the language pair has been tested by more than one formal MI test as with Turkish-Azerbaijani, we should simply average the results of all of the tests together, and this would be the latest word from science on the subject.

While it may be technically true that even the results of formal MI testing cannot completely prove that they have quantified MI between two languages perfectly, it makes sense to see this research as a sort of "as good as it gets" sort of conclusion.

It is difficult to quantify findings with precise numbers in the social sciences, as social science data tends to be messy with a lot of scatter and skew. Just because this is the inevitable result of testing endlessly variable human subjects does not mean that social scientists should always fall back on a nihilistic conclusion about questions in social sciences as is so often the case. Even if social science testing cannot be completely accurate in its conclusions, it is good to show the results with the best methodology we have as an "as good as it gets for now" conclusion.

The paper relied heavily on the findings of one study (Golubovic and Gooskens 2015a).

This study used three different types of tests to test both oral and written intelligibility of six different Slavic languages (Golubovic and Gooskens 2015a). Of the three tests, only the cloze tests tracked anywhere near close to native speaker MI estimates. Therefore, only the cloze figures from this study were used for oral and written intelligibility figures, and the results from their other tests were rejected.

However, the co-author of the study emphasized that her findings should be taken with a large grain of salt and should not to be relied upon as the last word on the MI of the languages in question: (Golubovic 2015).

“Nowhere do I claim that I have definitively measured the intelligibility of the six languages in question, only that the scores on a specific experiment with specific participants were such and such. What is the point of this, you might wonder? The point is simply recognizing that different tests and different participants may give different numbers. The numbers may vary, but since the same method was employed everywhere, the overall pattern should be the same.” (Golubovic 2015).

Results

64 different results on Slavic MI were obtained via formal intelligibility studies.

Of the oral intelligibility figures in under Results, 37 out of 241 (14%) were derived from formal intelligibility studies.

For the written intelligibility figures under Results, 30 out of 78 (38%) were derived from formal intelligibility studies.

Out of 219 oral and written intelligibility figures under Results, 64 (20%) were derived from formal intelligibility studies.

The languages that were tested scientifically have an asterisk after their name.

The other figures are from the variety of sources described above.

Figures

Slovene: *Oral intelligibility:* 100% of Standard Carinthian Slovene, Krapina Kaikavian, and Zagorje Kaikavian, 96% of Virje-Djurdjevac Kaikavian, 95% of Medjimurje Kaikavian, 90% of Standard Burgenland Croatian, 85% of Kaikavian, 80% of Serbo-Croatian* and

Karlovac Kaikavian, 55% of Čičarija Čakavian, 50% of Podchetrtek Kozjanski Kaikavian, 47% of Čakavian, 25% of Banat Bulgarian, 15% of Russian, 20% of Macedonian, 18% of Bulgarian* and Slovak*, 16% of Czech*, 12% of Polish* and Resian, 10% of Prekmurjian; 3% of Upper Sorbian, and 2% of Bednjanski Kaikavian. *Written intelligibility*: 94% of Serbo-Croatian*, 70% of Macedonian, 47% of Russian, 36% of Slovak*, 29% of Czech*, 22% of Bulgarian*, and 16% of Polish*.

Styrian Slovene: *Oral intelligibility*: 75% of Upper and Lower Carniolan Slovene.

Resian: *Oral intelligibility*: 0% of Slovene.

Kaikavian: *Oral intelligibility*: 90% of Prekmurjian, 80% of Bednjanski Kaikavian, 65% of West Gorski Kotar Kaikavian, 55% of Slovene, 41% of Čakavian (range 0-82%), and 0% of Dolinci Burgenland Croatian.

Serbo-Croatian: *Oral intelligibility*: 97% of Standard Burgenland Croatian, 96% of Serbian**, 95% of Bosnian**, and Croatian**, 93% of Montenegrin**, 81% of Šokački Štokavian; 86% of Virje-Djurdjevac Kaikavian, 80% of Karlovac Kaikavian, 75% of Ugljan Island Čakavian, 66% of Torlakian, 57% of Macedonian (range 25-90%), 55% of Pannonian Rusyn, 53% of Virovitica-Podravina Kaikavian, 44% of Slovene*, 37% of Kaikavian (range 3-86%) and Čakavian (range 0-75% of Čakavian), 27% of Czech*, 24% of Slovak*, 22% of Ukrainian, 20% of Bulgarian and Medjimurje Kaikavian *, 17% intelligibility of Belarussian and Dubravica Kaikavian, 15% of Rusyn Rusyn, 15% of Prekmurjian, 13% of West Gorski Kotar Kaikavian, 12% of Polish* and Molise Slavic, 10% of Upper Sorbian, 9% of Russian, 8% of Kashubian, 4% of Rhodope Mountains Pomak, 3% of Zagorje Kaikavian and Podravina Kaikavian, 1% of Resian and Bednjanski Kaikavian, and 0% of Dolinci Burgenland Croatian and Archaic Islander Čakavian. *Written intelligibility*: 93% of Šokački Štokavian, 77% of Macedonian, 64% of Slovene*, 45% of Belarussian, 40% of Czech* and Ukrainian, 35% of Slovak* and Rusyn, 30% of Kaikavian, 22% of Bulgarian*, 17% of Russian, 16% of Polish*, 12% of Kashubian, and 9% of Upper and Lower Sorbian.

Standard Croatian: *Oral intelligibility*: 99% of Purgeri Štokavian; 95% of Serbian, 93% of Bosnian, 92% of Montenegrin, 88% of Virje-Djurdjevac Kaikavian, 81% of Šokački Štokavian, 75% of Ugljan

Island Čakavian, 60% of Torlakian, 46% of Virovitica-Podravina Kaikavian, 37% of Čakavian (range 0-75%), 30% of Kaikavian (range 3-88%), 25% of Macedonian and Medjimurje Kaikavian, 20% of Belarussian and Ukrainian, 17% of Dubravica Kaikavian, 15% of Prekmurjian, 13% of West Gorski Kotar Kaikavian, 12% of Molise Croatian, 10% of Russian, 8% of Kashubian, 3% of Bednjanski Kaikavian, Podravina Kaikavian, and Zagorje Kaikavian, 1% of Resian, and 0% of Archaic Islander Čakavian. *Written intelligibility*: 93% of Šokački Štokavian, 45% of Belarussian, 40% of Ukrainian, 35% of Rusyn, 30% of Kaikavian, 14% of Russian, and 12% of Kashubian.

Čakavian: *Oral intelligibility*: 82% of Kaikavian, 45% of Russian, 40% of Slovene, Slovak, and Bednjanski Kaikavian, 32% of Archaic Islander Čakavian, 29% of Rhodope Mountains Pomak*, 20% of Bulgarian, and 5% of Molise Croatian.

Northern Čakavian: *Oral intelligibility*: 95% of Southern Čakavian.

Southern Čakavian: *Oral intelligibility*: 95% of Northern Čakavian and Kvarner Archipelago Middle Čakavian.

Archaic Islander Čakavian: *Oral intelligibility*: 98% of Old Čakavian†* and Old Church Slavonic†* and 71% of Rhodope Mountains Pomak*.

Molise Croatian: *Oral intelligibility*: 12% of Serbo-Croatian and 5% of Čakavian.

Standard Bosnian: *Oral intelligibility*: 100% of Serbian, and 5% of Medjimurje Kaikavian.

Standard Serbian: *Oral intelligibility*: 99% of Bosnian and Montenegrin, 97% of Croatian, 84% of Virje-Djurdjevac Kaikavian, 77% of Torlakian, 72% intelligibility of Kaikavian (range 60-84%), 60% of Virovitica-Podravina Kaikavian, 57% of Macedonian (range 25% in Vojvodina, 91% in Nis), 55% of Pannonian Rusyn, 30% of Bulgarian, 12% of Russian, 10% of Čakavian, 7% of Ukrainian, 0% of Bednjanski Kaikavian and Rhodope Mountains Pomak.

Torlakian: *Oral intelligibility*: 90% of Macedonian, 77% of Serbian, 30% of Bulgarian, 20% of Slovene and Russian, 15% of Kaikavian, 12% of Czech and Slovak, and 5% of Polish. *Written intelligibility*: 75% of Russian and 40% of Serbo-Croatian.

Macedonian: *Oral intelligibility*: 91% of Torlakian, 70% of Serbo-

Croatian, 65% of Bulgarian, 43% of Czech, 36% of Slovak, 27% of Polish, 24% of Slovene, 18% of Russian, 13% of Ukrainian, and 12% of Kashubian. *Written intelligibility*: 90% of Serbo-Croatian, 79% of Czech, 75% of Bulgarian, 73% of Slovene, 67% of Slovak, 41% of Russian, 40% of Polish, and 27% of Ukrainian.

Bulgarian: *Oral intelligibility*: 82% of Macedonian, 80% of Banat Bulgarian, 78% of Rup Bulgarian, 75% of Kostur-Korča Macedonian, 73% of Zlatograd Rup Pomak, 60% of Serbian, 50% of Croatian, 46% of Russian (varies from 7-85%), 40% of Torlakian, 35% of Belarussian and Ukrainian, 27% of Serbo-Croatian*, 20% of Slovene* and Čakavian, 13% of Slovak*, 10% of Greek Pomak, 10% of Czech* and Polish*, and 6% of Kashubian, 3% of Rhodope Mountains Pomak, and 0% of Upper Sorbian. *Written intelligibility*: 82% of Macedonian, 80% of Ukrainian, 71% of Russian, 32% of Serbo-Croatian*, 25% of Slovak*, 21% of Czech* and Slovene*, and 18% of Polish*.

Eastern Bulgarian: *Oral intelligibility*: 35% of Serbian, 30% of Serbo-Croatian, 25% of Croatian, and 12% of Western Macedonian.

Western Bulgarian: *Oral intelligibility*: 82% of Macedonian and 80% of Serbo-Croatian.

Slovak: *Oral intelligibility*: 99% of Pannonian Extraslovakian Slovak, 95% of Gemer Central Slovak, 94% of Czech*, 81% of Eastern Slovak, 55% of Pannonian Rusyn, 50% of Polish*, 45% of Rusyn, 40% of Western Ukrainian and Čakavian, 32% of Kashubian, 28% of Serbo-Croatian* and Old Church Slavonic, 27% of Belarussian, 25% of Croatian, 20% of Serbian and Ukrainian, 16% of Russian, 14% of Slovene*, 11% of Bulgarian*, and 5% of Macedonian. *Written intelligibility*: 98% of Czech*, 80% of Sulkovian Silesian, 75% of Saris Eastern Slovak, 70% of Silesian, 59% of Polish* and Upper Sorbian, 37% of Serbo-Croatian*, 35% of Russian and Ukrainian, 29% of Slovene*, 25% of Bulgarian*, and 0% of Macedonian.

Eastern Slovak: *Oral intelligibility*: 50% of Rusyn, 45% of Ukrainian, 40% of Polish and Kaikavian, and 30% of Russian.

Pannonian Rusyn: *Oral intelligibility*: 55% of Serbo-Croatian.

Czech: *Oral intelligibility*: 91% of Slovak*, 80% of Gemer Central Slovak and Pemci Czech, 70% of Pannonian Rusyn, 68% of Eastern Slovak, 52% of Upper Sorbian, 50% of Pannonian Extraslovakian Slovak and Serbian, 37% of Cieszyn Silesian, 36% of Polish*, 32% of

Rusyn, 27% of Ukrainian, 18% of Slovene* and Serbo-Croatian*, 15% of Bosnian, 13% of Russian, Bulgarian*, and Macedonian, 8% of Kashubian, and 5% of Belarussian. *Written intelligibility*: 100% of Upper Sorbian, 97% of Slovak*, 75% of Russian, 46% of Polish*, 45% of Serbo-Croatian*, 27% of Slovene*, and 19% of Bulgarian*.

Upper Sorbian: *Oral intelligibility*: 82% of Polabian†, 80% of Lower Sorbian, 33% of Czech*, 31% of Slovak*, and 8% of Polish*.

Lower Sorbian: *Oral intelligibility*: 80% of Upper Sorbian.

Polish: *Oral intelligibility*: 94% of Silesian Polish, 80% of Mazovian Polish, 77% of Lemko Rusyn, 75% of Rusyn, 70% of Slovio, 63% of Podhale Polish, 46% of Western Ukrainian, 41% of Slovak*, less than 41% of Kashubian*, 36% of Czech*, 32% of Belorussian, 30% of Ukrainian, 22% of Silesian (with no knowledge of Old Polish or German, much higher with that knowledge), 20% of Kostur-Korča Macedonian, 19% of Serbo-Croatian*, 15% of Slovene*, 14% of Upper Sorbian, 13% of Bulgarian* and Macedonian, 12% of Lower Sorbian, and 7% of Russian. *Written intelligibility*: 67% of Belarussian and Ukrainian, 60% of Russian, 54% of Slovak*, 50% of Czech*, 36% of Bulgarian*, 33% of Serbo-Croatian*, and 26% of Slovene*.

North Kashubian: *Oral intelligibility*: 85% of Pomeranian† and Slovincian† and 80% of South Kashubian.

South Kashubian: *Oral intelligibility*: 85% of Pomeranian† and Slovincian† and 80% of North Kashubian.

Cieszyn Silesian: *Oral intelligibility*: 35% of Czech.

Belarussian: *Oral intelligibility*: 80% of Ukrainian, 41% of Polish, and 25% of Slovak. *Written intelligibility*: 80% of Russian and 67% of Polish.

Podlachian: *Oral intelligibility*: 75% of Ukrainian.

Ukrainian: *Oral intelligibility*: 90% of Balachka, 85% of Eastern Ukrainian and Surzyk, 75% of Belarussian and Rusyn, 72% of Podlachian, 70% of, 67% of Lemko Rusyn, 50% of Russian, 45% of Eastern Slovak, 39% of Hutsul (range 7-75%), 35% of Polish and Russian, 25% of Kashubian, and 20% of Slovak and Serbo-Croatian. *Written intelligibility*: 90% of Slovak, 85% of Russian, 60% of Bulgarian, and 50% of Polish.

Balachka: *Oral intelligibility*: 100% of Ukrainian.

Canadian Ukrainian: *Oral intelligibility:* 5% of Russian.

Western Ukrainian: *Oral intelligibility:* 79% of Lemko Rusyn, 75% of Rusyn, 60% of Eastern Ukrainian, 57% of Polish and Eastern Slovak, 40% of Slovak, and 30% of Russian.

Eastern Ukrainian: *Oral intelligibility:* 85% of Ukrainian, 70% of Russian, and 60% of Western Ukrainian.

Surzyk: *Oral intelligibility:* 85% of Russian and Ukrainian.

Rusyn: *Oral intelligibility:* 78% of Ukrainian, 75% of Polish and Russian, and 50% of Eastern Slovak.

Lemko Rusyn: *Oral intelligibility:* 70% of Ukrainian.

Russian: *Oral intelligibility:* 85% of Surzyk and Ninilchik Russian, 75% of Belarussian and Rusyn, 70% of East Ukrainian, 60% of Balachka, 50% of Volgograd Russian, 45% of Eastern Slovak and Čakavian, 42% of Slovak, 40% of Ukrainian and Banat Bulgarian, 37% of Bulgarian, 35% of Croatian, 32% of Pannonian Rusyn and Western Ukrainian, 30% of Serbo-Croatian, 27% of Macedonian, 25% of Polish, 18% of Serbian, 17% of Lower and Upper Sorbian, 10% of Slovene, 8% of Kashubian, and 4% of Czech. *Written intelligibility:* 90% of Bulgarian, 85% of Belarussian and Ukrainian, 70% of Czech and Polish, 63% of Slovak, 50% of Serbo-Croatian and 25% of Slovene.

Ninilchik Russian: *Oral intelligibility:* 65% of Russian.

*Result from a formal intelligibility study

†An extinct language.

Overview of the Slavic Languages, Conclusions, and Discussion of the Study

South Slavic

There are different ways of looking at the South Slavic languages. One way is to divide them into West and East, with the Western

languages being the macrolanguages Slovene, Kaikavian, Štokavian, Čakavian. Slovene and Kaikavian seem to be part of one system, and Štokavian and Čakavian seem to be part of another system.

At any rate, Kaikavian is much closer to Slovene than it is to Štokavian (Kapović 2017).

Kaikavian was removed from Štokavian and put it in Slovene in the classification below.

The Eastern languages in this case would be Torlakian, Macedonian and the macrolanguage Bulgarian.

You can also divide these into Northern and Southern South Slavic along the same lines.

This division has the interesting results of severing Torlakian from Serbian, which it is considered to be a dialect of. However, Torlakian is closer to Macedonian than it is to Serbian, and it is more part of the Macedonian-Bulgarian system than part of the northern or western system.

In the Macedonian-Bulgarian system, it is much more a part of Macedonian than a part of Bulgarian, and Macedonian is closer to Torlakian than to Bulgarian, though it is close to both. However, in this classification, Torlakian was left as part of Serbian rather than move it into Macedonian, which seemed to extreme of a move. Further research and the evolution of linguistic consensus is needed to make such an extreme move.

This treatment would also move Serbian, Croatian, Montenegrin, and Bosnian into North or Western South Slavic, despite the somewhat different origins of these dialects. Although Torlakian is sometimes called Old Serbian, and it is actually a dialect of Štokavian, it seems better to move it out of Serbian and Štokavian altogether and into the Macedonian-Bulgarian system.

Čakavian is extremely ancient with its oldest forms going to back to the pre-Slavic era. One wonders exactly what connection it has with Štokavian. Yet there doesn't seem to be anywhere else to put it than with Štokavian. Nevertheless, placing both of them in Northern or Western South Slavic seems proper.

That Kaikavian and Čakavian are separate languages outside of Štokavian is beyond controversy in the realm of non-politicized

linguistics. Unfortunately, Croatian linguistics remains pathetically politicized, so politics continues to trump linguistic consensus and science here (Kapović 2017).

That Prekmurjian and Resian are separate languages outside of Slovene seems obvious to any observer. One wonders why more linguists have not stated the obvious, but breaking paradigms is a radical move in linguistics, and few wish to be the fall guy.

There seems to be an increasing consensus that Torlakian is more something that looks like a language as opposed to something that looks like a dialect. There is even a recognition in a few places that it is best placed outside of Serbian and into the southern or eastern group.

Slovene

Slovene is actually a macrolanguage consisting of at least 10 separate languages, Slovene, Resian, Prekmurjian, Kaikavian, Bednjanski Kaikavian, West Gorski Kotar Kaikavian, Strawberry Hill Gorski-Kotar Kaikavian, Gail Valley Slovene, Upper Carniolan Slovene, and Lower Pohorje Slovene. Standard Slovene is spoken by two million people in Slovenia.

Slovene, along with Slovak, is probably the most conservative Slavic language, although this is somewhat controversial. Case endings of both Slovene and Old Church Slavonic (OCS) look similar. Slovene did not participate in most of the Slavic innovations after OCS which can be seen in, for instance, Serbo-Croatian.

Slovene is the only large Indo-European language that has retained the dual number, although there are Slovene dialects that lack it. It also has the odd supine case, unique in Slavic.

In addition, it has retained much of its original OCS vocabulary. Keep in mind that OCS at its core is basically a South Slavic language.

An example of the conservatism of Slovene would be the Freising Manuscripts, some of the earliest records of the Slovene language and written over 1,000 years ago. They can still be read quite easily by Slovenes today, which shows that Slovene, like Icelandic, has changed little through the years.

These scripts are important for Slavistic research because they are one of the oldest extant Slavic documents (Kortlandt 2003).

In fact, the *Kozjak Styrian Slovene* dialect of the small town of Sveti Duh na Ostrem Vrhu in Northeastern Slovenia on the border with Austria looks remarkably like the language of these manuscripts, although the actual language of the scripts was spoken a bit to the north in Austria around Graz (Greenberg 2000, p.178).

In this part of Austria, some Slovene speakers remain to this day, a part of the small community of 12,000 Slovene speakers in Austria.

The language of the area from Northern Slovene to Western Slovak retains characteristics of these ancient scripts (Kortlandt 2003).

Some Slovenes have published pieces in which they are worried that Slovene might go extinct because almost all academic publication is done in English. Despite this, Slovene does not seem to be in danger of going extinct at the moment.

Slovene once had an imperfect and an aorist, but they have both gone out, with the imperfect first and then the aorist being replaced by the perfect.

Slovene has [47 different dialects which are 75-95% mutually intelligible](#). This diversity of dialects is surprising given how small the country is. There are dialects where neuter gender is absent, others that lack the dual, others than use masculine forms for females, others that are transitional to West Slavic (Slovene is South Slavic), some that still use the aorist tenses (Resian), yet more with odd / > w/u and g > h changes, some that form the future tense completely differently, and still more that have merged the feminine gender in some of the cases.

Many things have caused the diversity of these dialects.

Slovene is full of Germanic borrowings because it was under the influence of German for most of its history. Traditionally, Slovene was the low language of the commoners, while German was the language of high culture.

Written Slovene was based on the language of the Slovene Bible from 1583. This remained the standard until Slovene was standardized in 1850's. It would be as if all American English prose was written in the language of the King James Bible until 1850 and then suddenly

switched over to the vernacular.

Slovene underwent two periods of Serbo-Croatianization, one from 1860-1890 and another from 1920-1939. In the latter period, Slovene was officially a type of Serbo-Croatian. After World War 2, Slovenia became part of Yugoslavia, and Slovene became the official language, but the Yugoslavian military continued to use Serbo-Croatian.

Slovene has also been also under pressure from Hungarian, German, and Italian, which effected the language in the border regions.

The main dialects are divided into Upper Carniolan, Lower Carniolan, Styrian, *Pannonian*, Carinthian, Littoral, *Mixed Kojevce*, and Rovte dialects.

Standard Slovene was codified [based](#) on the *Gorenjska Eastern Upper Carniolan* and the *Deljeniska Lower Carniolan*, and to a much lesser extent on the Littoral dialects. Most of the other Carniolan features were removed during the formation of the standard language, which was based on the urban dialect of Ljubljana, during 1840-1860, so now the Carniolan dialects differ greatly from Standard Slovene. But an agreement was reached to base the vowel system on Upper Carniolan and the consonant system on Lower Carniolan.

The differences between the major Slovene dialects are so great that MI is often impaired, especially if they are from different regions (Sussex and Cubberley 2011). When communication between two dialect speakers is hindered, speakers use Standard Slovene. There is [more than one language present](#) in the Slovene dialects.

Lower Carniolan Slovene or Deljeniska is also one of the eight major dialects. It is another Slovene dialect that is regarded as hard to understand, this time [by speakers of Styrian Slovene](#). It is spoken in the south-central part of the country. Styrian Slovene speakers may have 75% of Lower Carniolan Slovene.

Rovte Slovene is one of the eight major Slovene dialects. It is spoken in mountainous West-Central Slovenia and is bordered by the Upper and Lower Carniolan and Littoral groups. Rovte dialects can be quite diverse. The group is bounded by the towns of Škofja Loka, Tolmin and Vrhnika.

Škofja Loka Slovene is a Rovte dialect spoken around the town of that name located 16 miles northeast of the capital Ljubljana. It resembles Upper Carniolan, especially the Selca subdialect. Speakers of

Standard Slovenian say they have a [hard time understanding this dialect](#). Intelligibility figures are not available.

Cernko Slovene is another Rovte dialect that differs quite a bit from the standard language, yet it is [still intelligible](#) with the rest of Slovene. However, some Slovenes from around Ljubljana say they have a hard time understanding it.

Carinthian Slovene is one of the main eight Slovene dialects. This dialect group is spoken in the far north of the country in the Alps up to and across the border into Austria in Slovene Corinthia and into Italy in a few villages in Udine Province. This seems to be one of the [hardest Slovene dialects of all](#) for the average Slovene to understand, harder than even some difficult Rovte and Styrian dialects. However, Slovenes can understand the Carinthian Slovene standard language [perfectly](#).

Littoral Slovene is one of the eight main dialect groups of Slovene. These dialects are spoken in the far west of Slovenia and over the border into Italy. Littoral dialects are some of the most diverse dialects in Slovenia. Some of the Littoral dialects spoken mostly in Italy but also in a couple of towns in Slovenia such as Torre Valley, Natisone Valley, Iudrio Valley, and Resian are actually separate languages.

Istrian Slovene is a Littoral dialect spoken in Slovene Istria. At least the dialect spoken around Sečovlje Salina Nature Park near the Croatian border is [almost impossible](#) for other Slovenes to understand. However, it is probably intelligible with Buzet. This includes the dialects of Piran and Koper. There is a large Italian speaking population here. In fact there are still towns where everyone speaks Italian.

The Italians here formally spoke Venetian and they spoke a different Venetian dialect in every town. However, Venetian speakers are now switching over to Italian. The local cuisine here also is very Venetian. This dialect has an Italian sound, feel, and rhythm to it along with many Italian borrowings. Many Croatians have migrated here for work in the last half century. Some of the older ones still speak Serbo-Croatian. However, their children all speak Slovene. This Croatian migration has left its mark on this dialect, which has some Croatian borrowings. While the dual is still widely used in Slovene, it has gone out in Istrian, replaced with the plural.

Šavrini Hills or Šhavrinski Slovene is an Istrian dialect. Other Slovenes say it is [almost impossible](#) to understand.

It is spoken from the Croatian-Slovene border north to a line running from Koper to Zazid. Speakers are referred to as Šhavrini or Brežani. The dialect is little known (Kalsbeek 1998).

Notranjski, Brkinski, or Dekani Slovene is spoken to the east of the Šavrini Hills dialect. This is the Brkinski subdialect of the Notranjski dialect. Speakers are called Brkini or Kraševski. It is also spoken alongside Čakavian along the Croatian-Slovene border (Kalsbeek 1998).

Styrian, Shtajerska, Pomjan, Istrski, or Stajerski Slovene is one of the eight main Slovene dialects. It is spoken in Slovene Styria and along the Lower and Central Sava Valleys. Styrian Slovene is traditionally seen as a Slovene-Kaikavian transitional dialect (M. Jembrih 2015). Styrian is full of Germanisms, even in the basic vocabulary. Some Slovenes have a hard time understanding Styrian.

Kozje-Bizeljsko Slovene is a Styrian dialect spoken along the Sava River on the border with Croatia. This is the southernmost Styrian dialect.

Podčetrtek Kozjanski Slovene, a Kozje-Bizeljsko subdialect, is actually a Kaikavian dialect. Slovenes have poor intelligibility of this dialect (M. Jembrih 2015).

Styrian Slovene intelligibility of other Slavic lects:

Styrian Slovene has 75% of Upper and Lower Carniolan Slovene.

Slovene intelligibility of other Slavic lects:

The study of Slovene MI was eased by the presence of many Slovene informants. There was only one personal informant, and they were only slightly useful.

Slovenes have full intelligibility of Standard Carinthian Slovene, [Krapina Kaikavian](#), and Zagorje Kaikavian at 100%. Krapina is spoken near Podčetrtek, and Podčetrtek is Kaikavian-Slovene transitional.

Slovenes have full intelligibility of 96% of Virje-Djurdjevac Kaikavian, a Kaikavian-Štokavian transitional dialect.

Slovenes have [full understanding](#) of Standard Burgenland Croatian at 90%.

Slovenes have very high intelligibility of Kaikavian at 85%. This makes sense as Kaikavian is actually a part of Macro-Slovene, not Macro-Serbo-Croatian.

Slovene intelligibility of Serbo-Croatian is very high, 80%* (Golubovic and Gooskens 2015a). Written is full at 94%* (Golubovic and Gooskens 2015a). There is a strong confounding factor of bilingual learning here. For instance, 70% of Slovenes have some knowledge of Serbo-Croatian. Slovene adults can often understand Serbo-Croatian very well, as Serbo-Croatian was the official language of Yugoslavia. Also most Slovene adults grew up watching Serbo-Croatian TV.

However these same people report that young Slovenes, especially children, say they cannot understand Serbo-Croatian TV at all. The Slovene written language used to contain more words similar to Serbo-Croatian, but most of those words are now considered outdated and archaic and are no longer used. Most of those words went out long before independence in 1991. One problem is that when Slovenes speak Serbo-Croatian nowadays, the vocabulary they use is often 30 years old from Yugoslavian days.

Slovene has high 80% intelligibility of Karlovac Kaikavian.

Some Slovene lects have poor understanding of other Slovene lects. For instance, Styrian Slovene has only 75% intelligibility of Upper Carinthian Slovene, adding weight to the case that Upper Carinthian is a separate language as it was classified in this treatment.

Slovenes have 55% of Čičarija Čakavian, a very strange Čakavian dialect spoken right near the Slovenian border that may be transitional to Slovene.

Slovene intelligibility of Čakavian and Kaikavian is good at 47%. Slovenes understand Čakavian and Kaikavian better than Serbo-Croatians do. More than half of Slovene is cognate with Čakavian, Kaikavian, and Old Štokavian, which is nearly extinct. A lot of these are old words, and older Slovenes would probably know more of them than younger Slovenes. Also some Slovene dialects will have more of these old words than Standard Slovene. This is the best explanation for the fairly high Slovene intelligibility of Čakavian.

Even basic vocabulary is very similar between Slovene, Kaikavian, and Čakavian (Petrov 2015).

Standard Slovene intelligibility of Podčetrtek Kozjanski Slovene,

traditionally seen as a Slovene dialect transitional to Kaikavian but actually a dialect of Kaikavian, is not full. Slovene has 50% of Podchetrtek Kozjanski Kaikavian (M. Jembrih 2019). What is odd is that Slovene has full intelligibility of Krapina Kaikavian but only 50% intelligibility of Podchetrtek Kaikavian, while the two are right next to each other and both are transitional to Slovene.

Slovene intelligibility of Banat Bulgarian is a bit higher than that of Bulgarian at 25%.

Slovenes understand 21% of spoken Russian. Written understanding is higher at 40%.

Slovene intelligibility of Macedonian is poor at 20%. Written intelligibility of Macedonian is dramatically higher at 70%. The low oral figure makes sense if we consider Northern and Southern South Slavic to be dramatically different systems. The much higher written factor may be explained by the resemblance of Macedonian written language to Serbo-Croatian, which Slovenes have full understanding of.

Slovene oral intelligibility of Slovak is much lower at 18%* (Golubovic and Gooskens 2015a). Written is adequate at 36%* (Golubovic and Gooskens 2015a).

Slovene understanding of spoken Bulgarian is the same as with Slovak, 18%* (Golubovic and Gooskens 2015a). Slovenes understand 22% of written Bulgarian* (Golubovic and Gooskens 2015a).

Slovene oral intelligibility of Czech is low, at 16%* (Golubovic and Gooskens 2015a). Written is higher, 29%* (Golubovic and Gooskens 2015a).

A Slovene would need years of exposure to speak Czech or Russian well.

Slovene has dramatically low 12% intelligibility of Resian.

Slovenes understand spoken Polish poorly, 12%* (Golubovic and Gooskens 2015a). Slovene understanding of written Polish is also very low at 17%* (Golubovic and Gooskens 2015a).

Slovenes have [very poor intelligibility](#) of Prekmurjian at 10%. This is fascinating as Prekmurjian is formally seen as a Slovene dialect. When in truth it is actually a part of Kaikavian instead.

Slovene has very poor understanding of Upper Sorbian at 3%.

Slovenes cannot understand Bednjanski at all, being nearly unintelligible at 2%.

When it comes to Bednjanski Kaikavian, Slovenes are like everyone else. It appears that only their fellow Kaikavians along with their neighbors the Čakavian have decent understanding of Bednjanski, while everyone else is left shaking their heads in incomprehension. Kaikavian's decent understanding of Bednjanski is due to Bednjanski being an outlying Kaikavian lect. Čakavian's decent understanding may be because Čakavian is a Croatian language with ancient roots, as the first Croatian language was Old Čakavian. Bednjanski's roots also go back to this same ancient Old Croatian language.

Upper Carniolan Slovene or Gorenjsko is a Slovene dialect, one of the eight major ones, spoken in the center-northwest of the country. Standard Slovenian and Styrian Slovene speakers find it hard to understand, the Styrian speakers [in particular](#). Upper Carniolan is regarded along with Prekmurjian and Resian as one of the hardest dialects to understand in Slovenia.

All other things being equal, it does appear that Upper Carniolan is a separate language. Slovenes from the Styrian area to the capital seem to have a hard time with it, and Slovenes generally rank it on the same level as Prekmurjian, which is definitely a separate language. Styrian may have 75% intelligibility of Upper Carniolan, adding weight to the notion that this may be a separate language.

Zilsko or Gail Valley Slovene is a Carinthian lect spoken in the Alps of Southern Carinthia in Austria, Northeastern Udine in Italy and Northeastern Upper Carniola in Slovenia. It has particularly difficult intelligibility with the rest of Slovene and Slovenes say it is [as hard to understand as Prekmurjian or Resian](#).

South Pohorje Slovene seems to be the most difficult Styrian dialect to understand. It is spoken around Maribor. A good case can be made that it is a separate language, as most Slovenes [have a hard time understanding it](#).

Resian

Resian is an extremely divergent Slovene lect, typically characterized as a Slovene dialect, part of the Littoral dialect group, but the truth is that it is a completely separate language. The Resian language is composed of four dialects: Resian Proper, Torre Valley, Natisone Valley, and Iudrio Valley. Torre Valley and Natisone Valley are formally characterized as Littoral Slovene dialects along with Resian. Iudrio Valley is usually not listed among the formally described Slovene dialects, but perhaps it should be.

In this sense, Resian Proper is a dialect of the Resian language. Resian has essentially zero intelligibility with the rest of Slovene. Slovene speakers cannot even get the general context of Resian speech as they can with Prekmurjian, so Resian is even farther from Standard Slovene than Prekmurjian.

Resian is spoken by Slovenes living outside of Slovenia in Italy in the Resian Valley near the Slovene border.

This part of Italy is near Trieste, the Venetian-speaking area, and even closer to the Friulian-speaking area in far northeastern Italy. The entire Resian-speaking area is part of an area called called Slavia Friuliana.

Resian is so different from Slovene that it even has its own separate orthography, which suggests strongly that it is a separate language and not a Slovene dialect. Resian's orthography, grammar, phonetics, and vocabulary are all dramatically different from Slovene. Resian Proper is quite a bit more archaic than even Torre Valley Slovene. The Resian language has many Slavic archaisms.

While the aorist has gone out in all of Slovene, it has been retained in Resian, but only in a few verbs (Comrie and Corbett 1998).

Resian may have been split from the rest of Slavic for a long time as Slavic speakers began settling in Slavia Friuliana, migrating from Southern Poland through Slovenia to far northeast Italy in the Proto-Slavic period in the year 720. At some point, they were called "Resians" = "Russians" by Slovenes who encountered them probably because their lect was so far from South Slavic.

Slavia Friuliana came under control of the Duchy of Friuli in the 800's. In 1077, control was transferred to Patriarch of Aquileia and the region was known as Patria de Friuli, so it was still under Friulian control, so the area was under Friulian control for 600 years.

The region was conquered by Venice in 1420, and was under Venetian control until 1797. During this period, the region was called Venetian Schiavonia or "Venetian Slavland." Hence there were 377 years of Venetian influence.

There has traditionally been no education in Resian in these valleys, so most Resians never become fluent in Slovene. Many Resians cannot understand Standard Slovene at all. Others with access to Slovene TV and radio have a better understanding of Slovene due to bilingual learning. However, in the early 1990's, a local school began offering Slovene classes.

Friulian is heavily spoken in Slavia Friuliana and has already replaced Resian as the primary language of communication in the villages of Montenars, Tarcento, Nimis, Attimis, Torreano, and Prepotto.

Torre Valley Resian is spoken mostly in Italy but also in a few towns in Slovenia and is [unintelligible](#) to most Slovenes. It is very archaic, heavily Romanized and has many Friulian words. This dialect is spoken in the towns of Tapiana and Lusevera in Italy and in Breginj and Livek in Slovenia. It is now threatened and has only a very reduced number of speakers.

Natisone Valley Resian is a dialect of the Resian language spoken in San Pietro al Natisone, San Leonardo, Torreano, Pulfero, Stregna, Drenchia, and Savogna.

Iudro Valley Resian is also a dialect of the Resian language.

Torre Valley, Natisone Valley, and Iudro Valley are all very closely related. Resian Proper is more distant to but nevertheless intelligible with the other three. 19th Century linguists who studied these lects felt that the four Resian lects were all separate Slavic languages. However, given their full MI, it is best to lump them into a separate Slavic language called Resian as opposed to four separate languages.

Resian intelligibility of other Slavic lects:

Study of Resian MI was near impossible due to a complete absence of Resian informants.

Resian intelligibility of Slovenian is 0%.

Kaikavian

Kaikavian is a separate Slovene language within Macro-Slovene. It is usually considered to be a dialect of Serbo-Croatian. More progressively, it is now seen as a separate language within Macro-Serbo-Croatian. But the truth is even stranger because actually it is completely outside of Serbo-Croatian and a part of Slovene itself spoken in Northwest Croatia and is similar to Slovene. And even odder than that is that Kaikavian itself is a macrolanguage consisting of four languages, Kaikavian, Prekmurjian, Gorski-Kotar Kaikavian, and Strawberry Hill Gorski-Kotar Kaikavian.

It is formally characterized along with Čakavian as a Serbo-Croatian dialect, but many Croatian linguists regard Kaikavian as a bonafide language, and this is the best characterization of Kaikavian (Silić 1998).

In this treatment, I decided to make the radical decision of moving Kaikavian out of Serbo-Croatian and moving it to Slovene.

Kaikavian has a particularly close relationship to Slovene.

In fact, a good argument can be made that Kaikavian is closer to Slovene than it is to Serbo-Croatian. In fact, if Kaikavian had ended up in Slovenia instead of Croatia, there's no doubt that it would be considered a Slovene lect and not a Croatian lect (Kapović 2017).

If Kaikavian is indeed closer to Slovene than to Serbo-Croatian, then it would be part of Macro-Slovene instead of part of Macro-Croatian. Since Kaikavian is having a very hard time even getting recognized as a separate language by the ISO, mostly due to resistance by nationalistic Croatian linguists, the question is better set aside for now until it is discussed further and a consensus can be arrived at.

That Kaikavian and Čakavian are separate languages from Standard Croatian is quite clear. For one thing, the differences between Čakavian, Kaikavian, and Standard Croatian are greater than the differences between the Slovene dialects (Kapović 2017), and the Slovene dialects are quite divergent.

Of course, Kaikavian is not a dialect of anything, certainly not Serbo-Croatian and not even of Slovene. The best solution seems to be to classify Kaikavian not as a Slovene dialect but as a separate language

within the rubric of Macro-Slovene or the Slovene macrolanguage.

This same treatment almost moves Prekmurjian out of Slovene as a Slovene dialect, as it is nothing of the sort, and even Slovenes refer to it as a language and not a dialect; and at any rate, it has much more in common with Kaikavian than with Slovene. The best solution here is to make Prekmurjian part of Macro-Kaikavian. Despite Prekmurjian's long and illustrious history, Kaikavian seems to have more gravitas, so Prekmurjian should be a part of Kaikavian and not the other way around.

Kaikavian was proven to be a full language complete with its own internal dialects in the 1930's (Ivšić 1936). Kaikavian is actually a macrolanguage consisting of four languages as described above.

A change request was submitted to SIL to recognize Kaikavian as a separate language in 2013, but it was rejected on technical, not linguistic grounds. In their rejection, SIL acknowledged that Kaikavian was a separate language but said that there was no way to fit the code into the present ISO-3 code naming system for Croatia.

Kaikavian split from the rest of South Slavic at the same time Čakavian, Slovene, and Štokavian did, that is when Western South Slavic began to break apart in the 900's (Matasovic 2008). Since its split 1,100 years ago, Kaikavian has pursued a different evolutionary trajectory than the rest of Western South Slavic.

In the 1500's, Kaikavian began to be developed in a standard literary form. Kaikavian was the standard literary language of Croatia from the 1500's to 1850. It was also taught in schools during this time. From the 1500's to 1900, a large corpus of Kaikavian literature was written (Lončarić 1996).

In truth, this literary language was not Kaikavian per se as spoken today but another type of Kaikavian called Ikavian-Kaikavian. This was the official public language of the Kingdom of Croatia outside of Dalmatia from the 1400's-1600's under the rule of princes such as Zrinski and Frangipani (M. Jembrih 2016).

In the last half of the 19th Century, Croatians adopted the new Serbo-Croatian standard that was created for Croatia, Serbia, Bosnia, and Montenegro.

Linguists Ljudevit Gaj from Croatia and Vuk Karadžić from Serbia got together in 1850 and crafted a new language, choosing Štokavian as

the language of choice and basing their Štokavian on the Ijekavian Štokavian dialect of Eastern Herzegovina. The particular Ijekavian dialect was chosen because Karadžić came from an Ijekavian-Štokavian-speaking town which has since Ekavianized. The process was sealed by an document called Vienna Literary Agreement.

In creating the literary language, the two scholars based the new language almost exclusively on Štokavian. Almost no features or words of Čakavian or Kaikavian were used in creating this modern Serbo-Croatian. With the signing of this document, the decline in Kaikavian began in earnest.

Kaikavian was removed from public use after 1900, hence writing in the standard Kaikavian literary language was curtailed (Fishman and Garcia 2011). As Kaikavian was replaced by Štokavian as the official language of Croatia, many Kaikavians either stopped speaking their language or forgot a lot of the old words.

Nevertheless, writing continues in various Kaikavian dialects which still retain some connection to the old literary language, although some of the old lexicon and grammar are going out (M. Jembrih 2014).

Although Croatian linguists recognized Kaikavian as a separate language as long ago as the 1930's (Ivšić 1936), any suggestions that Kaikavian is a separate language are censored on Croatian TV (M. Jembrih 2014). However, there is an official TV station run by the Croatian government, TV-Varaždin, broadcast from Mt. Vranilec. This program, incredibly enough, broadcasts nearly 24-7. There are also Kaikavian schools, mostly in Zagorje, and there is even a special Department of Kaikavian Studies in the town of Čakovec, a satellite of Zagreb University.

However Josep Silić, a very prominent Croatian linguist, has recently authored a work in which he states that Kaikavian is a full Slavic language and not a dialect of Serbo-Croatian. This is important because Silić is one of the men who sits on the language committee involved in the ongoing standardization of the official Croatian language (Silić 1998).

The ISO has recently accepted a proposal from the Kaikavian Renaissance Association to list the Kaikavian literary language written from the 1500's-1900 as a recognized language with an ISO code of

kjv. The literary language itself is no longer written, but works written in it are still used in public, for instance in dramas and church masses (M. Jembrih 2014).

This is heartening, although Kaikavian as an existing spoken language also needs to be recognized as a living language instead of a dialect of Croatian.

Kaikavian differs from the other Slavic languages spoken in Croatia in that it has many Hungarian and German loans (M. Jembrih 2014). There are quite a few German loans and a lesser number of loans from Hungarian and Latin. Many of the German loans are found in neither Croatian nor Slovene.

Kaikavian is definitely closer to Slovene than it is to Čakavian (M. Jembrih 2014), and it has a particularly close relationship with Prekmurjian Slovene (Univerza v Mariboru in Zgodovinsko Društvo Maribor 1988, Lončarić 2007), which is actually a completely separate language from Slovene Proper. Kaikavian is even further from Štokavian than it is to Čakavian, and it is closer to Slovene than to either of them.

For example:

“house”

Slovene: *hiša*

Kajkavian: *hiža*

Neo-Štokavian (*i, ije, or e yat*): *kuća*

In fact the Kaikavian-speaking region used to be referred to as Kaikavian Slavonia or Slovenje. Even the name of the region – Slovenje – looks like the word Slovenia. The Kaikavian-Slovene region tended to be more connected towards Catholic Central Europe and less towards the Balkan region as a whole (M. Jembrih 2014).

Kaikavian and Slovene share the use of the word *kai* for the interrogative pronoun “what”, whereas Čakavian does not use this term. Most Slovenes also use *kai*, but some far to the north around Kranj and in Austria in the Kärnten region, where some Slovene speakers remain to this day, do not use *kai* (M. Jembrih 2014).

The Kärnten was originally Slovene speaking, but it became Germanized after Carantanians from the Kärnten area and the Carnioles of the Kranj and Kaikavian regions, under the command of

the Kaikavian Prince Ljudevit, lost a war to the Carolingians (M. Jembrih 2014).

Kaikavian also has a strong but little noticed relationship with Slovak, which makes sense if it is close to Slovene, as Slovak and Slovene are thought to have ancient connections, and Slovak has many South Slavic features, which is why it is often listed as the "central" Slavic language. Kaikavian is much closer to Slovak structurally than it is to Čakavian and Štokavian. Past tense ends in *-l* in both Slovak and Kaikavian and verb building and syntax is the same, while syntax is slightly different in Štokavian. In addition, both languages were originally called Slovenski.

Presently Kaikavian is coming under deep influence from Standard Croatian and is becoming greatly Štokavized. A lot of old Kaikavian words are being replaced with their Croatian equivalents.

Kaikavian does have one interesting feature. Early in its development it developed *ü* out of *u*. Later, under Štokavian influence it changed back to *u* again. Traces of the old *ü* can be found in archaic Kaikavian lects such as Bednjanski, Podravje, and Turopolje. It is presently vanishing from the latter two dialects (M. Jembrih 2019).

It has also been retained in the Prekmurjian language and the Buzet and *Prlekija* dialects of Slovene. Prlekija is a dialect of the Pannonian Group, and it has full MI with Prekmurjian. The Prekmurjian language is also part of Pannonian Slovene. It is also present in a number of eastern dialects of Slovene. This shows that these three Slovene lects had an ancient connection to Kaikavian (M. Jembrih 2019).

The real hard Kaikavian basilect that is very hard to understand is not spoken much anymore. On the other hand, a mesolect which is combination of Kaikavian and Štokavian is very commonly spoken by residents of Northern and Northeastern Croatia. It is even heard of the floor of the Croatian Parliament.

Kaikavian is actually spoken by 31% of Croatians, following closely the 56% who speak Štokavian.

There are three types of Kaikavian spoken in Croatia, Ikavian-Ekavian Kaikavian, Ekavian Kaikavian, and Ijekavian Kaikavian.

Ijekavian Kaikavian is the most numerous dialect, encompassing 18% of all Croats and 60% of all Kaikavian speakers. This dialect includes all Kaikavian speech except for that spoken in the Zagorje Hills and

Ikavian-Ekavian Kaikavian regions. The center of this dialect is in Zagreb.

A hard Kaikavian with many old words still exists in some areas, mostly in rural districts. For instance, in the Zagorje Hills and on the Medjimurje Plains, there are [many people](#) who can barely speak Standard Croatian. This more pure Kaikavian is spoken only in the historical districts of large cities.

Northern Kaikavian is the name of the Kaikavian spoken in Northern and Northeastern Croatia (Kapović 2017). It encompasses many of the subdialects below.

There was formerly a dialect continuum between Kaikavian and Old Štokavian similar to the Čakavian-Old Štokavian continuum discovered by Peco. This old transition zone has been completely interrupted, and the boundary between Kaikavian and Štokavian in Croatia today is a sharp one.

Podravje or Podravina Kaikavian, part of the *Virovitica-Podravina* dialect, the furthest east Kaikavian dialect. It is fully intelligible with East Gorski-Kotar, the furthest west dialect (M. Jembrih 2019). It borders on Old Štokavian-speaking Slavonia, so it is a Kaikavian dialect transitional to Štokavian. This dialect is definitely hard Kaikavian. It is also very archaic.

Podravina just means the Drava River Basin, but in this case it probably refers to the southern part of Virovitica-Podravina County, with Virovitica in the north and Podravina in the south. The southern border of Virovitica-Podravina County is probably the southeastern end of the Kaikavian zone in Croatia. South and east of that is all Slavonian Old Štokavian.

People from this region are called Podravec, a little known ethnonym (M. Jembrih 2019).

Virovitica-Podravina Kaikavian is a dialect similar to Virje-Djurdjevac that is spoken in Slavonia in Virovitica-Podravina County eight miles west of the Hungarian border 18 miles south of the Virje-Đurđevac region. This is a Kaikavian-Štokavian transitional dialect. Serbo-Croatian has better intelligibility of this dialect than with other Kaikavian dialects but not as good as with Virje-Djurdjevac. This dialect is quite Štokavianized but not as much as Virje-Djurdjevac. These Slavonian dialects of Kaikavian seem to be moderately to

heavily Štokavized, and some of them are also Slovenized. The German influence on Virovitica-Podravina is also quite strong.

There is still the remains of the old dialect continuum between Slovene and Kaikavian.

Kopravina Križevci Kaikavian is a dialect spoken in the county by that name. There are at least eight different major Kaikavian dialects in this county, and the whole county speaks Kaikavian, albeit a type that is heavily-admixed with Slavonian Old Štokavian.

Virje-Đurđevac Kaikavian is a Kaikavian-Štokavian transitional dialect, one of the last remains of the old Kaikavian-Štokavian transition area. It is spoken around the towns of Virje and Đurđevac in Northeastern Croatia five miles west of the Hungarian border. It is made up of about 2/3 Kaikavian words and 1/3 Štokavian words. Štokavian and Slovene both have near-full intelligibility with this dialect. Virje-Đurđevac is a subdialect of Kopravina Križevci.

Križevci-Podravina Kaikavian is the name for a group of Kaikavian dialects spoken in Slavonia to the south of the Medjimurje Plain along the Hungarian border and east to near Zagreb. It encompasses all of Kopravina Križevci County, including Virje-Đurđevac Kaikavian spoken in the southeast of the county over to Virovitica-Podravina County at the southeastern edge of the Kaikavian zone.

Dialects in this area are characterized by a shift of tone from circumflex to neoacute: *posêkel* > *posěkel*. Here you see that the *e* shifts from a long fall tone a long acute tone (Ivšić 1911). This change, also known as the rise of the neocircumflex, only occurs in certain Kaikavian dialects that are in contact with Slovene along the Slovene border, as the change initiated in Slovene (Kapović 2017).

Medjimurje Kaikavian, spoken on the Medjimurje Plain 44 miles northeast of Zagreb on the borders of Slovenia and Hungary, is one of these hard Kaikavian dialects. This dialect still has fluent speakers in their 20's. This is one of the strongholds of the Kaikavian language, and it is still common to find people here who barely speak Serbo-Croatian.

Turopolje-Posavina Kaikavian is the name of a dialect spoken in the Sava River Basin on extending 5-6 miles on both sides of the river southeast of Zagreb around the Turopolje area in the west to Ivanić Grad in the east. Posavina is a name for the Sava River Basin. In this

dialect, the long rise and long fall accents change to short accents and move backwards or retract from the second syllable to the first. For example: *pos ê kel* > *p ò Sekel*. Here the neocircumflex in the second syllable shortened and retracted to a short acute accent in the first syllable (Ivšić 1911).

Ivanić Grad Kaikavian is a Kaikavian dialect that still has speakers of all ages. Children come to school as Kaikavian monolinguals, and everyone in the town speaks Kaikavian. Ivanić Grad is located 20 miles southeast of Zagreb, so Kaikavian is spoken a ways southeast of the capital in addition to the dialects extending far to the northeast (Pavličević-Franić and Aladrović 2017). This is part of the Posavina dialect. This particular part of Posavina on the east side of the Sava is called Posavina Bregi or Moslavina (Celinič 2010).

Sveti Ivan Zelina Kaikavian is another dialect that is spoken by all ages in the town. As in Ivanić Grad, children come to school as Kaikavian monolinguals and learn Standard Croatian as a second language at school. Sveti Ivan Zelina is located in Zagreb County four miles south of the Zagorje Hills and 16 miles northeast of Zagreb (Pavličević-Franić and Aladrović 2017). If children are still coming to school as Kaikavian monolinguals, then Kaikavian will probably survive at least until the end of the 20th Century.

Turopolje Kaikavian is spoken in the Turopolje region a bit southeast of Zagreb between Zagreb and Sisak.

It is quite archaic (M. Jembrih 2019). The villages of Horvati and Zdenčina still speak Ikavian-Ekavian Kaikavian (Vermeer 1982).

The area around Zagreb has always been the center of the Kaikavian region.

Zagreb Kaikavian or Agramer Kaikavian is a hard Kaikavian dialect spoken in the old city of in Central Zagreb, and even there it is spoken only by small population of mostly old people in the suburb of Varaždin and the Agramer District of Zagreb (Bujan 2015). This is the real Zagreb Kaikavian. In 2011, there were still 14,000 speakers of this hard Agramer Kaikavian, which resembles the Kaikavian spoken in the 19th Century. This shows that 1.75% of the population of Zagreb still speaks this old pure Kaikavian.

A dictionary of this language was published in 2012. It ran to over 1,000 pages and included over 20,000 words.

Speakers of Agramer are called *Agrámeri* or "*laškovuličánski*."

Most of what passes as Kaikavian spoken in Zagreb is actually now so Štokavized that it is no longer Kaikavian is now a Štokavian dialect called Purgeri, a Kaikavianized Štokavian dialect.

However, there are other reports that Agramer is still common in the capital, where it is said that if a Serbo-Croatian speaker start walking in any direction from the city center, within five minutes you will hear someone who is unintelligible.

There are 1.45 million Kaikavian speakers around Zagreb, 410,000 Purgeri speakers and over 1 million speakers of hard Kaikavian.

These Kaikavian speakers in the Zagreb region now have 12 separate dictionaries of their dialects. A new grammar has also been written.

The areas of newer Štokavian-speaking immigrants are mostly in the new suburbs built in the 20th Century in the eastern Dubrava district, the southern Novi Zagreb district, and in the suburb of Sisek. There are three major full-time radio stations in Zagreb which can be heard from Southern Slovenia to Northern Bosnia. *Radio Zagreb* is mostly Shtokavian, *Radio Sljeme* is in Purgeri Shtokavian, and *Radio KAJ* is completely in Kaikavian.

All of Kaikavian is mutually intelligible across its entire range except for Bednjanski and West Gorski-Kotar.

Dubravinca Kaikavian is spoken just south of the Zagorje Hills. Since it is not Zagorje Kaikavian, it cannot be Ekavian. It is spoken near the famous Lower Sutla River dialects now seen as Ikavian-Ekavian, but it is not a part of Lower Sutla. Instead this seems to be like East Gorski-Kotar, one of the last Ijekavian Kaikavian dialects spoken west of Zagreb, the rest having been influenced so much by Middle Čakavian that they have all gone over to Ikavian-Ekavian under its influence.

Dubravinca was used in a popular Croatian TV show, hopefully with subtitles, as Croatian intelligibility of this dialect is as poor as with the rest of hard Kaikavian. Kaikavian is still heavily spoken in this town.

East Gorski-Kotar Kaikavian is the furthest west Kaikavian dialect. This is the dialect of the Kaikavian poet Ivan Goran Kovačić, who lived in Lukov Dol, a town in this region. Lukov Dol was just outside the Medieval Kaikavian Bishopric, but it still speaks Kaikavian (M. Jembrih 2019). It has full MI with Podravina, the furthest east Kaikavian

dialect. On the other side of the Kupa River in Slovenia, locals speak the same language as the Croats speak. So East Gorski-Kotar is spoken [over the border for a ways into Slovenia also](#).

Ekavian Kaikavian is spoken only in the Zagorje Hills. It is spoken by 10% of all Croats and 30% of all Kaikavian speakers.

Zagorje Kaikavian, spoken near Zagreb in the Zagorje Hills 7-26 miles north of Zagreb, is another of the hard Kaikavian dialects. This dialect still has middle-aged fluent speakers. Zagorje retains a pronoun that looks like the remains of the old dual: *midva* "we two". Slovene also retains the old dual from Proto-Slavic. The dual was still a part of Proto-Slavic in the 800's (Nuorluoto 2010).

Medjimurje and Zagorje Kaikavian have a close relationship to Slovene. Medjimurje is close to some Pannonian Slovene dialects, [in particular](#) the Prekmurjian language (Lončarić 2007).

Zagorje is [close](#) to some Styrian Slovene dialects spoken across the border which are usually seen as a Slovene-Kaikavian transitional. Slovene may have full intelligibility of Zagorje, but so does the rest of Kaikavian, complicating classification. It seems best for now to leave it in Kaikavian, as the Kaikavian elements outweigh the Slovene factor. Zagorje is [very archaic](#), and it is one of the hardest Kaikavian dialects for Croats to understand (Petričević 2019). It is common to find people in this region who can hardly speak Serbo-Croatian.

Upper Sutla Kaikavian or Gornjosutlanski is a Ekavian Kaikavian dialect spoken along the Upper Sutla River near the border with Slovenia from Hum na Sutli to the headwaters a bit to the east. This dialect is interesting because it is the only Kaikavian dialect that did not undergo certain changes along with Čakavian and Western Štokavian, namely a retracted accent in the infinitive, *l*-participles, and present verbal adverbs ending in *-ći*. These changes were conditioned by neither phonology nor prosody as in the Neo-Štokavian and Slovene accent retractions. Instead, the change was purely morphological (Kapović 2017).

Hum na Sutli Kaikavian is an Ekavian Kaikavian dialect that is specifically protected by the Croatian state. This is an Upper Sutla Kaikavian dialect. It is spoken right on the Slovene border. This dialect resembles Slovene more than any other Kaikavian dialect. This was actually the Yugoslavian Communist leader Marshall Tito's native

dialect and he still retained a Kaikavian accent when he spoke Serbo-Croatian.

Stajerski Kaikavian, an Ekavian Kaikavian dialect spoken in Slovenia along the Croatian border around Podčetrtek across from the Croatian border towns of Hum na Sutli seven miles to the north and Kumrovec eight miles to the south, is now said to be a Slovene dialect transitional to Kaikavian, part of the *Kozje-Bizeljvsko* subdialect of *Styrian Slovene*. However, it is not intelligible with Slovene, so it cannot possibly be a Slovene dialect. This is part of the Zagorje dialect.

Podčetrtek Kozjanski Kaikavian is a Stajerski Ekavian Kaikavian dialect traditionally known as a Slovene dialect called *Podčetrtek Kozjanski or Govorokoli Podčetrteka Slovene*, is [not intelligible at all](#) with *Standard or Kraj Slovene*. We know it is Kaikavian because it is 80% Kaikavian and only 20% Slovene. However, it is intelligible with Kaikavian spoken across the border in Croatia, hence it is best seen as a Kaikavian dialect transitional to Slovene (M. Jembrih 2015).

It has been described as 80% Kaikavian (often Old Kaikavian) and 20% Slovene (M. Jembrih 2015, A. Jembrih 2003). It is [particularly close](#) to Zagorje Kaikavian. Podčetrtek is in Slovenia 27 miles north of Zagreb on the Croatian-Slovenian border.

Ikavian-Ekavian Kaikavian, spoken by 3% of Croats or 10% of Kaikavians, is the remains of the Kaikavian-Čakavian transitional dialect continuum in Croatia, which has now nearly collapsed. However, some of the old Kaikavian-Čakavian transitional dialects are still spoken.

They are often classed as "Ikavian Kaikavian", which probably does not exist, or even as Čakavian. Actually they are Kaikavian, not Čakavian, and they are Ikavian-Ekavian Kaikavian according to Jakubinskij's Law, which generally also applies to Middle Čakavian (Mihaljević 2002). The change to Ikavian-Ekavian probably came from Middle Čakavian influence in the transition zone.

There is a major problem with the theory that these dialects are Ikavian Kaikavian that no one has ever mentioned. Where did this Ikavian Kaikavian come from? The theory is that Ijekavian Kaikavian changed to Ikavian Kaikavian in the transition zone with Čakavian. So the Ikavian must have transferred from Ikavian Čakavian to

Kaikavian.

But there is a huge problem here.

Ikavian Čakavian dialects are spoken in two Čakavian groups – Southwestern Istrian Čakavian and Southern Čakavian. Let's look at where they are spoken. SW Istrian Čakavian is spoken in the south and west of Istria. They moved there from Dalmatia long ago, which has always spoken Ikavian, either Ikavian Čakavian, Ikavian Čakavian-Štokavian, or now mostly Ikavian Štokavian. SW Istrian Čakavian has only been in Southwestern Istria for 300 years and it's never been anywhere else. But Istria nowhere near the proposed contact zone which extends far inland from Istria. An Istrian dialect could not possibly have been the Čakavian spoken in the Kaikavian-Čakavian contact zone.

What about their source dialect? They came from an Ikavian-speaking region in Dalmatia, but that is too far south to be a source of the Čakavian in the contact zone. Yes, this zone did spread eastward into the Dalmatian Highlands to the northern Bosnian border near Binac. But these were Štokavian dialects, not Čakavian ones. So SW Istrian Čakavian either from Istria or its source population in Dalmatia could not possibly have been the source of the Čakavian in the contact zone.

The only other Ikavian Čakavian is Southern Čakavian. Southern Čakavian has the advantage of having once extended far inland in Croatia according to studies about the homelands of the Burgenland Croatian Southern Čakavians who fled their homeland in the 1500's. But even there, this theorized zone only went as far north as the Una River. That's far below the Kaikavian-Čakavian dialect continuum.

If Čakavian did indeed used to extend further east and is the source of the "Ikavian" dialects, it would have had to have been Middle Čakavian. If you draw a line from the Middle Čakavian area eastwards, you will find that some Middle Čakavian dialects line up perfectly with the Čakavian-Kaikavian contact zone. So Middle Čakavian must have been the source of the "Ikavian," and indeed there is evidence that it did extend far to the east, as it still has some isolated dialects further inland quite close to this contact zone.

There are also Middle Čakavian dialects stranded far to the east beside the Sava River along its northern bank in Croatia at the Vrbas

River Estuary on the border with Northern Bosnia and Southeastern Croatia (Slavonia) in the villages of Davor, Orubica, Siće, and Magićmala. These dialects have been there since 1500, and they look remarkably like Kvarner Archipelago Middle Čakavian. So Middle Čakavian once extended as far east as Western Slavonia. If you go north from there, you end up in Vitrovica. Yet the transition zone for whatever reason only extends to the Lower Sutla River a bit northeast of Zagreb.

So Middle Čakavian once extended far enough north and east to have been the source of contact zone Čakavian. But there is a problem with the Ikavian theory, and that is that Middle Čakavian has never been Ikavian. Instead it is Ikavian-Ekavian according to Jakubinskij's Law.

And indeed, it turns out that a very large number of these "Ikavian" Kaikavian dialects have recently been reanalyzed as actually being Ikavian-Ekavian (Mihaljević 2002). We don't have proof that *all* the dialects said to be Ikavian are actually Ikavian-Ekavian, but a huge number of the other dialects just like them in this transition zone have been proven so (Mihaljević 2002), therefore they must be too.

And if indeed these holdouts are Ikavian, where did this Ikavian come from? It had to have come from Middle Čakavian, and it has never been Ikavian, only Ikavian-Ekavian.

All of this leads to the inescapable conclusion that all of the "Ikavian" Kaikavian dialects are actually Ikavian-Ekavian due to contact with Ikavian-Ekavian Middle Čakavian. Therefore, Ikavian Kaikavian does not exist and never did. The only source for it in Čakavian is too far to the south to be the source of the contact zone.

This will probably be controversial, and it might make people mad, but I believe that there is no other possible conclusion, and further research on the few dialects not yet proven to be Ikavian-Ekavian Kaikavian will, or actually must, show that that is indeed what they are.

Ikavian-Ekavian Kaikavian is spoken in the far west near the coast in Hreljin, Fužine, Lokve, Sleme and Zagorje inland from Rijeka; around Ogulin in Ogulin, Oštarija, and Mihaljevići; in Karlovac; in the Bregan River Valley in the Zumberak Hills; in Zdihovo, Vukova Gorica, Mrežnica, Duga Resa, Jelsa, Severin na Kupi Ribnik, Ozalj, and Pribić in the Pokupje Valley in the basin of the Kupa River (Mihaljević 2002).

It is also spoken in Marija Gorica, Brdovec, and Donja Pušća (Vermeer 1982) in the Southwestern Zagorje Hills on the east bank of the Lower Sutla River; and in Horvati, Gornja Zdenčina, and Donja Zdenčina along the Mrežnica River in the Turopolje region (Mihaljević 2002) and in Sredičko a bit to the southeast.

Ogulin township is 5-10 miles west of the Mrežnica River and 15 miles south of Karlovac. The Žumberak Hills extend in a 20 mile band east to west 10-25 miles north of Karlovac along the Croatian-Slovenian border. The Pokupje Valley extends north of Karlovac 10 miles to Slovenian border.

The speech of Fužine, Lokve, Sleme, Ogulin is fairly archaic, similar that of the Lower Sutla Valley. In the Pokupje Valley dialects, there are much fewer archaic words, most having been replaced by typical new Balkanisms, with few archaic words among them.

In Ogulin township, Tounj to the east speaks Čakavian, and Zagorje to the west speaks Kaikavian. The villages of Ogulin, Oštarija, and Mihaljevići to the north speak dialects that are difficult to characterize.

Here is an example of a word list from what a linguist calls a Čakavian dialect in the villages of Oštarije, Mihaljevići, and Zagorje in the old Čakavian-Kaikavian dialect continuum zone away from the islands and the Adriatic Sea. This was published in a Čakavian magazine as an example of an odd Čakavian dialect (M. Jembrih 2014).

The word collector of the list above referred to this dialect as Čakavian only because it has *ča* instead of *kai* for the interrogative pronoun "what". However, the mere presence of *ča* or *kai* alone in a dialect is probably not sufficient to characterize it as either Čakavian or Kaikavian (M. Jembrih 2014).

Yet this "Čakavian" dialect also has many Kaikavian features such as German loans, prosthetic *v-* before *u*: *vučitelj* "teacher" and *vučenik* "student" (M. Jembrih 2014). Prosthetic *v-* is also found in a number of Slovene dialects but is absent in others.

This is a classic example of Kaikavian intrusion into a Čakavian dialect. The *v-* is probably an old retention from Common Slavic, whereas **v-* went to *u-* in Štokavian (Matasović 2008).

Therefore, a better analysis is that none of these are Čakavian dialects. Instead, Oštarija, Ogulin, and Mihaljevići speak Kaikavian dialects transitional to Čakavian or for shorthand Ikavian-Ekavian

Kaikavian.

Nevertheless, Ikavian-Ekavian Kaikavian looks like the remains of the old Kaikavian-Čakavian dialect transitional dialect extending along the border from Istria all the way east to north of Karlovac.

Prigorje Kaikavian, an Ikavian-Ekavian Kaikavian dialect spoken southwest of Zagreb, is clearly transitional to Čakavian (Kapović 2017). This seems to include Pokupje Valley dialects like Ozalj. Prigorje is [very archaic](#). Croats find Prigorje quite hard to understand (Petričević 2019).

Pokupje Valley Kaikavian is an Ikavian-Ekavian Kaikavian dialect (Mihaljević 2002) spoken in the Pokupje Valley or the basin of the Kupa River which forms the border between Croatia and Slovenia from East Gorski Kotar past the Črnomelj or Bela Krajina area north of Karlovac in down into Croatia by Ozalj towards Karlovac. All along this river basin which runs along the border, Pokupje Valley is also spoken [for a ways into Slovenia](#).

Karlovac Kaikavian is an Ikavian-Ekavian Kaikavian dialect (Mihaljević 2002) spoken near the city of Karlovac by the old Kaikavian-Čakavian transition zone. Slovene has good but not full understanding of this dialect, which may be particularly close to Slovene. It is also spoken [across the border in Slovenia](#) for a ways.

Donjosutlanski or Lower Sutla Kaikavian is the name of an Ikavian-Ekavian Kaikavian dialect spoken around the Lower Sutla River (Kapović 2017) 10 miles west of downtown Zagreb. It was originally spoken in Southern Croatia near Bosnia.

Kaikavian intelligibility of other Slavic lects:

The study of Kaikavian MI was greatly helped by two Croatian informants, one a speaker of Kaikavian and a linguist. A linguistics professor was also helpful.

When tested, Kaikavian was found to have 90% intelligibility of Prekmurjian (M. Jembrih 2019). However, we only had one informant, and he was a linguist who knew five languages and had some knowledge of Slovene, so he's not your average Kaikavian speaker, who will probably score a bit lower than 90%. Nevertheless, sources are unanimous that Kaikavian has [good intelligibility](#) of the Prekmurjian language. Nevertheless, intelligibility is generally reported as partial, as speakers of the two languages have [a hard](#)

time understanding each other. One estimate put intelligibility at 82%, but that is only a rough guess.

Kaikavian has 65% intelligibility of Western Gorski-Kotar Kaikavian (M. Jembrih 2019).

Kaikavian intelligibility of Bednjanski Kaikavian ranges from 40-84%, with an average of 62% (M. Jembrih 2015).

Kaikavian has 60% understanding of Čakavian. Reports that Kaikavian and Čakavian are mutually intelligible are wrong. In truth, Kaikavian has variable intelligibility of Čakavian. In the border areas between Kaikavian and Čakavian zones, Kaikavian intelligibility of Čakavian is 80% or even higher. However, Kaikavian intelligibility of the hard Čakavian spoken around Split is much lower, well below 50% (M. Jembrih 2015).

Kaikavian intelligibility of Slovene is 55%. Although it is said that Kaikavian and Slovene are mutually intelligible, this is not correct. The truth is that Kaikavian is closer to Slovene than it is to either Serbo-Croatian or Čakavian (M. Jembrih 2015).

Nevertheless, Kaikavian speakers say that they do not understand much Slovene. Even Maribor Styrian Slovene, a South Pohorje dialect spoken in the city of Maribor that is supposedly a Slovene-Kaikavian transitional dialect, is understood poorly by Kaikavian speakers. It is quite interesting that Kaikavian has about as much intelligibility of the diverse Macro-Kaikavian languages as it does with Slovene. This shows how much these languages have diverged from Kaikavian Proper.

The Dolinci Burgenland Croatian Old Čakavian dialect spoken in Unterpullendorf cannot be understood at all by Kaikavian speakers (M. Jembrih 2015).

Bednjanski Kaikavian

Bednja, Baednjunski or Bednjanski Kaikavian is the most most divergent Kaikavian lect in Northern Croatia.

It has similarities with West Slavic languages such as Polish and Czech.

A theory is that they were the last Slavicized remnant of the Pannonian Gepides, a Gothic-speaking group. Both theories place Bednjanski Kaikavian in the general Pannonian region which has long connected Kaikavian and Czech/Slovak, especially Slovak, hence connecting West and South Slavic. The theory is that after Magyarization of Pannonia, groups of Slavic speakers moved south into the Slovenia/Croatia region.

The presence of Indo-Iranian words in this language implies that is in part the remains of the original Croats, thought to be an Indo-Iranian people from the Old Orient who were replaced by Slavic immigrants in the 600's. There are similar archaisms found in Archaic Islander Čakavian spoken along the Dalmatian coast and Istrian Peninsula.

It is now spoken in Varaždin County in a triangle bounded by the towns of Bednja, Gornje Jesenje, Trakošćan, including the villages of Šaša, Vrbno and Pleš, and smaller villages around them. This area is 28 miles north of Zagreb, five miles south of the Slovene border, and 3 miles east of the Zagorje Hills. The terrain is characterized by forest-covered hills. The Croatian Ministry of Culture recently voted to support the preservation of Bednjanski Kaikavian as a Croatian cultural monument.

This language was isolated geographically so it retained Old Kaikavian elements and innovated others on its own. In contrast to other Kaikavian lects, Bednjanski Kaikavian is 100% Kaikavian, with no Slovene or Čakavian words. It differs from the rest of Kaikavian primarily in phonology. For instance, it has a vowel that sounds like German *ü* (M. Jembrih 2015).

Kaikavian has high but not full intelligibility with the rest of Bednjanski Kaikavian (M. Jembrih 2015). Therefore, at least on an MI basis, Bednjanski Kaikavian qualifies as a separate language.

Bednjanski Kaikavian is nearly unintelligible to most Serbo-Croatian speakers.

West Gorski Kotar Kaikavian or Zapadnogoranski, usually considered to be the westernmost Kaikavian dialect, is actually a separate language because it is not intelligible with any other lect due to its archaic nature (M. Jembrih 2019). It is spoken in Northwestern Croatia on the Slovene border around Delnice, Crni Lug, Prezid, and

Čabar. It is also spoken for a ways across the border [for a ways into Slovenia](#) in the entire Gorski-Kotar region.

The latest scholarly thinking is that this is actually a dialect of Slovene transitional to Kaikavian (Kapović 2017). However, it is [not intelligible](#) with Slovene at all. Neither is it intelligible with Kaikavian or Čakavian (M. Jembrih 2019). The principal elements seem to be Slovene and Kaikavian, but the basic nature of the language seems to be Kaikavian.

Delnice is 20 miles west of Rijeka, and Čabar is 22 miles north of Rijeka on the Slovene border. Prezid is also on the border five miles northwest of Čabar. West Gorski-Kotar is the name of the Croatian province where this is spoken.

The portion of the language spoken around Čabar, Prezid, and the headwaters of the Čabranka River is ancient, as it is nonpalatalizing. It mostly has old speakers now. It also seems to be the remains of an ancient Čakavian-Kaikavian transition zone. This portion of the zone apparently became heavily Slovenized.

This area was sparsely populated in the 1100's, and the small population fled northwards across the Kupa River in the 1400's to escape the Turkish threat.

In the late 1600's after the threat waned, most of them all returned. However, it is thought that many Slovenes came with them, mostly from the Dolenjska/Bela Krajina area. Some of those who fled remained in Slovenia. Nobles in West Gorski-Kotar called for metallurgists and blacksmiths to come from the Kranjska region in Upper Carniola near the Italian and Austrian border, as a lot of mines had opened in that area. Besides Slovenes, many Ličani and Bunjevci from Dalmatia also arrived, so this added an Ikavian Štokavian element.

West Gorski-Kotar has a lot of Čakavian influences in addition to the heavy Slovene layer. It also has archaic Kaikavian influences (M. Jembrih 2019). The best theory is that this an old Kaikavian dialect transitional to Čakavian, part of the old Čakavian-Kaikavian transition zone, that has become significantly Slovenized. The main question is whether it is part of Macro-Kaikavian or Macro-Slovene. It seems best for the moment to go against expert opinion and take it out of Slovene and put it in Kaikavian, perhaps a Kaikavian language

transitional to Slovene.

The Medieval Kaikavian Bishopric of Zagreb went all the way up to Bosiljevo in West Gorski-Kotar, so this language was under direct influence of Kaikavian for a long time (M. Jembrih 2019). This is further evidence for its basic Kaikavian nature.

There are other clues to the Kaikavian nature of this language. West Gorski-Kotar uses an archaic long numbering system for enumeration found in Old Kaikavian, whereas Slovene uses the German numbering system such as "one and twenty (21)", and "two and twenty (22)" (M. Jembrih 2019).

The old long numbering system in West Gorski-Kotar looks like *dva na deset* "(12)", *tri na deset* "(13)", and *dvadeset dva* "(22)." Whereas Kaikavian, Čakavian, and Štokavian use *dvanajst* "(12)", *trinajst* "(13)", etc., but they all used to use the old long numbering system (M. Jembrih 2019).

Sesvete Kaikavian is a dialect of West Gorski-Kotar spoken in the eastern outskirts of Zagreb around Sesvete, where it is spoken only by the elderly.

Strawberry Hill Gorski-Kotar Kaikavian (Strawberry Hill for short) is an outlying form of Kaikavian spoken by Gorski-Kotar immigrants who left the area around 1900 to immigrate to Strawberry Hill, a neighborhood in Kansas City, Kansas. Speakers came from both East and West Gorski Kotar areas. They came from [Ribnik and Lukov Dol in East Gorski Kotar and Čabar in West Gorski-Kotar](#) (Filipovic 1997). Although the West and East Gorski-Kotar are parts of separate languages in Croatia, apparently they spoke a merged dialect of both East and West Gorski-Kotar which became intelligible to both groups, a sort of a Gorski-Kotar koine.

A large group of them came at once and settled in this area, which became a [large Croatian community](#) with everything one needed in terms of stores, etc. The men spoke Strawberry Hill among themselves and their American supervisors learned just enough Strawberry Hill to communicate with them (Filipovic 1997).

As they lived in an entirely Strawberry Hill-speaking neighborhood, there was no need to learn English, and the people of this generation [never learned](#) to speak English (Filipovic 1997). However, the 2nd, 3rd, and 4th generations apparently spoke both Strawberry Hill and

English.

Speakers of Strawberry Hill say that when they go back to their home villages in Croatia, they [find it hard to communicate](#) with locals speaking Gorski-Kotar dialects because the language has changed so much in 100 years (Filipovic 1997). Hence, the American form of Gorski-Kotar Kaikavian spoken in Kansas is a separate language.

When Filipovic conducted his research in Strawberry Hill in the 1970's, Strawberry Hill was still maintained through three generations via church, family, work, ethnic organizations, immigrant-owned businesses, St. John the Baptist Catholic Church and school (Filipovic 1997).

The original Strawberry Hill community had 20,000 speakers (Filipovic 1997).

But by 2012, there were only [300-400 Croatian-Americans](#) in Kansas City, and most of them did not speak Strawberry Hill. The generation of the first wave had died, and their children were mostly elderly. Most of this group no longer spoken Strawberry Hill. The generation of the second and third waves all spoke Strawberry Hill, and they had all passed it on to their children (Glasgow 2012).

Although Strawberry Hill was very much alive in the 1970's, it was best described as threatened and shifting by 2012. Nevertheless, in 2012 there were still child speakers of Strawberry Hill between the ages of 3-15 (Glasgow 2012), so it looks like the language will survive until at least the last half of the 21st Century.

Strawberry Hill Gorski-Kotar Kaikavian intelligibility of other Slavic lects:

There were no Strawberry Hill informants, making the study of this language difficult.

Strawberry Hill has difficult intelligibility of Gorski-Kotar in recent years when speakers go back to visit their hometowns as the languages have changed so much. This was the rationale for splitting off Strawberry Hill in the first place. Nevertheless, actual figures are not available.

Prekmurjjan

Prekmurjjan is typically characterized as one of the most famous Slovene dialects, and along with Resian and Gail Valley, is considered among the three most difficult dialects to understand in Slovenia. But even Slovenes often refer to it not as a Slovene dialect but as a separate language as a language. Every objective linguist who has looked at Prekmurjjan has determined that it is definitely a separate language form Slovene.

However, even they are not quite correct because Prekmurjjan isn't even a dialect of Slovene at all. It was very difficult for me to do this, but Prekmurjjan really has to be taken out of Slovene and put into the Kaikavian language because the two together really are a single language.

And I now believe after much deliberation and hand-wringing that Kaikavian needs to be taken out of Serbo-Croatian altogether because it is much closer to Slovene than it is to Štokavian, and it has completely different roots than Čakavian.

Further, they were connected with two different states. Čakavian was really the language of the Croatian state that existed in the Middle Ages. Croatia had different boundaries then. What is now Northern Croatia was part of the Kaikavian state, which lasted until 1700-1800. Kaikavian was the official language of this state and everyone spoke Kaikavian. A few hundred years ago, the Kaikavian state merged with the Croatian state to the south and modern Croatia was created.

Looking at the matter objectively, Kaikavian has a lot more in common with Slovene than it does with Čakavian. Čakavian doesn't have much of a connection with Štokavian either, and hard Čakavian is hardly more intelligible than hard Kaikavian. On the other hand, Čakavian isn't close to any other language either, unless you want to suggest Rhodope Mountains Pomak in Bulgaria.

But this connection is based more on Čakavian's being the original Old Croatian and Rhodope Mountains Pomak being so archaic, perhaps both being remnants of Proto-Mediterranean Slavic or Proto-South Slavic speech, perhaps a koine that was used by sailors sailing between the Black Sea to the Adriatic. But it hardly makes sense to put Pomak in Bulgaria on the Black Sea in with Čakavian all the way

over on the Adriatic.

Furthermore, when people think of the original Old Croatian, they think of Archaic Islander Čakavian because that was truly the original Croatian speech. Kaikavian was also present around that time or somewhat later, but only certain Kaikavian languages such as Gorski-Kotar and Bednjanski, which also have traces of the Old Croatian language in them. So while Čakavian is hardly connected with Štokavian, it's not really connected with any other local language either, and given its long history as Old Croatian and being the official language of the Medieval Croatian state, it makes sense to put it in Serbo-Croatian because there's nowhere else to put it.

Prekmurjians are Catholics of course, like their neighbors the Hungarians, Slovenes, and Croats. The Prekmurjian language was actually standardized long ago, and works have been written in Prekmurjian for hundreds of years. If you go to the region, there are signs in Prekmurjian everywhere and everyone speaks it.

So the solution seems to be making Prekmurjian a part of Kaikavian and then making Kaikavian a part of Macro-Slovene. Both decisions are extremely radical and will probably make a lot of people angry and cause a lot of outrage, but as I have considered this problem for years, I believe that this solution is scientifically motivated and justified no matter what popular opinion might be. At any rate, prominent Serbo-Croatian linguists have been suggesting that Kaikavian is closer to Slovene than Serbo-Croatian (Kapović 2017), and everyone knows that Prekmurjian has a particularly close relationship with Kaikavian.

Prekmurjian is still widely spoken though it is repressed and has no official status.

Kaikavian and Slovene share other things in common. They have similar customs, clothing, and festivals. Further, while Kaikavians say that they are certainly Croats, many of them also proudly say that they are Slovenes. Slovenes see Kaikavians as brothers and fellow Slovenes who happen to live in a neighboring country. It's clear that there is a lot of warmth and camaraderie between the two groups.

Prekmurjian is generally considered to be part of the Pannonian dialect group of Slovene. There are officially four major dialects in Pannonian and unofficially even more. Further, all Pannonian dialects

are mutually intelligible with each other and with Prekmurjian, so all of the Pannonian dialects should really be categorized as dialects of Prekmurjian.

One also wonders whether we should call it Prekmurjian Slovene owing to its present existence as a Slovene dialect. But even Slovenes recognize that Prekmurjian is not a Slovene dialect, and they typically refer to it as the Prekmurjian language. If we are going to take Prekmurjian out of Slovene and putting it into Macro-Kaikavian, we can hardly continue to call it a Slovene dialect.

Serbo-Croatian or Štokavian

Serbo-Croatian itself is best seen as a macrolanguage, with Štokavian or Standard Serbo-Croatian being only one of up to an incredible minimum of 13 different languages within this larger group. The eight languages would be *Neo-Štokavian*, *Dalmatian*, *Bunjevci*, *Čakavian*, *Gan-Veyān Archaic Islander Čakavian*, *Rab Island Archaic Islander Čakavian*, *Kyrška Bešeda Archaic Islander Čakavian*, *Susak Island Archaic Islander Čakavian*, *Komyzjanski Archaic Islander Čakavian*, *Korzulot Archaic Islander Čakavian*, *Molise Croatian*, *Burgenland Croatian*, and *Torlakian*.

MI among the Slavic languages of the Balkans in general is much exaggerated except for the official languages Bosnian, Croatian, Serbian, and Montenegrin, which have full MI.

Research into the Slavic lects of the former Yugoslavia has been problematic. Slavic research in the region during the Yugoslav era often served political rather than linguistic goals (M. Jembrih 2014). This trend continues to this day.

East Herzegovinan was chosen as the standard of Neo-Štokavian or Serbo-Croatian. This decision was made rather arbitrarily in the 1800's.

The main Neo-Štokavian dialects of *Croatian*, *Serbian*, *Montenegrin*, and *Bosnian*, now recognized as separate languages on political and sociological grounds and not the basis of linguistic science, are generally fully mutually intelligible to a high degree. All were based

predominantly on the Neo-Štokavian standard. However, the new languages have been standardized on a different basis in each country.

Pitch accent in Serbo-Croatian is highly complex, and 2nd language learners never seem to get it right, hence non-native speakers are easy to identify.

Not all Serbo-Croatian dialects have pitch accent. Many Serbian dialects do not have it. Belgrade Serbian and Novi Sad Serbian lack pitch accent, and *Nis Serbian* probably does also. Pitch accent was present in the official language because Serbo-Croatian was codified based on *Eastern Herzegovinan Bosnian Štokavian*, which has considerable pitch accent.

Proto-Slavic had free and mobile accent, and it has been retained in all daughter languages, only in different forms in its descendants. It evolved into pitch accent and vowel length in Slovene and Serbo-Croatian. In Russian, Belarussian, Ukrainian, Bulgarian, Macedonian, Czech, Slovak, Polish, and Sorbian, it evolved into stress position (Sussex and Cubberley 2011, p. 154).

Synthetic past tenses seem to be going out in Serbo-Croatian as they have in Slovene. The aorist and the imperfect have both become less commonly used with time. The imperfect is mostly used by older people. Nevertheless, the use of the imperfect remains quite scattered, and it is still common at least in some places or dialects or among certain age groups. However, the vast majority of the population no longer uses the imperfect, mostly because they can't distinguish it from the aorist.

The aorist is much more alive in Serbo-Croatian than the imperfect is, particularly in Croatia.

Štokavian (and Čakavian and to a much lesser extent Kaikavian for that matter) is divided into different types depending on the *yat* reflex. For instance, Ikavian *di, vrime, tribat, cili* become Ijekavian *gdje, vrijeme, trebat, cijeli*. Ijekavian is the standard for Serbo-Croatian, but Ikavian is also widely used in Croatia.

Štokavian is divided into two types, Old Štokavian and New Štokavian. The Bosnian, Serbian, Croatian, and Montenegrin standards are all based on New Štokavian.

Šćakavian is an Old Štokavian dialect spoken in Northern Serbia,

Eastern Croatia, Dalmatia, Bosnia, and parts of Herzegovina.

In Western Herzegovina, Šćakavian dialects with Čakavian influences are now present mostly along the Una River in Northern Bosnia along the Croatian border and along the Rama River and in Duvno Valley 17-25 miles inland from Dalmatia in an area called the Dalmatian Highlands. These dialects are heavily Čakavianized and are theorized to be the remains of an Old Štokavian-Čakavian transition zone. This old transition zone is discussed in more depth in the Čakavian section.

Novoštokavian or *New Štokavian* is the name for what we now know as Štokavian. It actually spread north and west with the Turks. There were other Štokavian dialects spoken at this time, but they were mostly pushed out by the spread of Novoštokavian. These other Štokavian dialects either went extinct or now exist only in scattered locations.

Many Croats feel that Ikavian should have been made the standard in Croatia, as it is so much more widely used (Petrov 2015).

Western Ijekavian Neo-Štokavian is widely spoken in Bosnia and Herzegovina, less so in Montenegro and Croatia, and to an even lesser extent in Serbia.

Serbo-Croatian (Štokavian) intelligibility of other Slavic lects:

Serbo-Croatian intelligibility theorizing was state of the art, with many excellent informants and a couple of great personal informants. Distortion due to nationalism was surprisingly almost zero. In this region, speakers either understand each other well or they don't, so there's not a lot of room for intellectual sophists to play. Even the "they are all dialects" crowd admits that a number of the "dialects" have poor MI. They base their 'dialects, not languages' arguments on things other than MI or even linguistics, frankly.

Serbs and Croats have [full intelligibility](#) of Standard Burgenland Croatian at 97% (M. Jembrih 2015). The official language was standardized around Standard Croatian.

Serbo-Croatian understanding of Serbian is full at 96%. Figure applies to non-Serbs.

Serbo-Croatian comprehension of Bosnian is full at 95%. The figure applies to non-Bosnians.

Serbo-Croatians understand Croatian fully at 95%. The percentage

applies to non-Croats.

Serbo-Croatian intelligibility of Montenegrin is full at 93%. The rating is for non-Montenegrins only.

Serbo-Croatian has marginal intelligibility of Virje-Djurdjevac Kaikavian of at most 86%. For many speakers even in Croatia, it may be less than that. This is a Kaikavian-Serbo-Croatian transitional dialect.

Serbo-Croatian understanding of Šokački Štokavian is marginal at 81% (Petričević 2019). However, written is full at 93%. The oral figure is good evidence that this indeed is a separate language.

Serbo-Croatian comprehension of Karlovac Kaikavian is very high at 80%.

Serbo-Croatians have a very high understanding of Ugljan Island Čakavian at 75% (Petričević 2019). This is Čakavian-Štokavian transitional.

Serbo-Croatians have good understanding of Torlakian at 66%. Croats have poorer understanding of this language than Serbs do. Serbian intelligibility of Torlakian is better than for Macedonian.

Serbo-Croatian 57% understanding of Macedonian. Written is very good at 77%. The oral figure ranges from 25-90%, but those figures apply to Serbians only. The 25% figure is from Vojvodina in the far north, while Nis Serbians who border Macedonia have much higher intelligibility of Macedonian at 90%. Serbo-Croatian speakers from the northern end of the speech area do not understand Macedonian well, but as one goes south in the zone, Macedonian becomes easier for them to understand.

The intelligibility figure is dragged down by Croatian and Northern Serbian poor intelligibility of Macedonian. Part of the problem between Serbo-Croatian and Macedonian is that so many of the basic words - *be, do, this, that, where* - are different, however much of the rest of the vocabulary is the same. Serbo-Croatian speakers can learn to understand Macedonian well after some exposure.

Serbo-Croatian has high intelligibility of Pannonian Rusyn at 55%. However, the sample size here was very small. The figure applies to Serbs only. This is actually an Eastern Slovak dialect. Pannonian Rusyn has undergone heavy influence from Serbo-Croatian, and the

added Serbo-Croatian vocabulary is the reason for the high intelligibility because Serbo-Croatian intelligibility of Slovak Proper is only 24%.

Serbo-Croatians have 53% understanding of Virovitica-Podravina Kaikavian. This is Kaikavian-Štokavian transitional.

Serbo-Croatian intelligibility of Slovene is 44%* (Golubovic and Gooskens 2015a). Written intelligibility is higher at 64%* (Golubovic and Gooskens 2015a). Some Serbs and Croats report good intelligibility of Styrian Slovene, but their understanding is probably not full, as even Kaikavian speakers understand little of Maribor Styrian Slovene.

Serbo-Croatian has 37% intelligibility of Kaikavian (Petričević 2019). Written understanding is actually worse at 30%, a strange figure as written is almost always higher than oral. Although a few Kaikavian dialects are well-understood by Serbo-Croatian speakers, most of the language is hardly intelligible to them. Reports that these languages are mutually intelligible are in error. Serbo-Croatian inherent intelligibility of hard Kaikavian, as with Čakavian, is very low, near zero.

For most Serbo-Croatian speakers, Kaikavian is simply a different language along the same lines as Slovene. Kaikavian vocabulary, grammar, and phonology differ a lot from Serbo-Croatian. Even the sound inventory is different. For instance, Kaikavian has at least nine vowels, whereas Serbo-Croatian has only five (Petrov 2015).

Serbo-Croatian has 37% intelligibility of Čakavian, the same as for Kaikavian. Intelligibility ranges from 0-75%, a very high skew. Serbo-Croatian inherent intelligibility of pure Čakavian, as with Slovene and Kaikavian, is very low, close to zero. Serbo-Croatian is not mutually intelligible with Čakavian as some reports claim. Serbo-Croatian intelligibility of Čakavian also differs according to dialect. Croats can understand the heavily-Štokavized lects such as Ugljan Island well.

But in the case of someone speaking the hard Čakavian on one of the islands and another person speaking pure Croatian from the Slavonia region of Croatia (the northeastern part of the country) such as Osijek Štokavian, the Slavonian Croatian speaker would have a very hard time understanding the Čakavian speaker, if they could understand them at all. Serbo-Croatian intelligibility of Čakavian is hampered by

the many Romance loans, archaic words, diphthongs, and terminal accent present in Čakavian (Kolanović 2015).

Serbo-Croatian intelligibility of spoken Czech is only 27%* (Golubovic and Gooskens 2015a). Written Czech is better understood, 40%* (Golubovic and Gooskens 2015a).

Serbo-Croatian has limited comprehension of spoken Slovak, 24%* (Golubovic and Gooskens 2015a). Written comprehension is better at 35%* (Golubovic and Gooskens 2015a).

Serbo-Croatian understands Ukrainian at 22%. It has 40% understanding of Ukrainian when it is written (Petričević 2019).

Serbo-Croatian intelligibility of Bulgarian speakers is quite low at only 20%* (Golubovic and Gooskens 2015a). Written is also quite poor at 22%* (Golubovic and Gooskens 2015a).

Serbo-Croatian understanding of Medjimurje Kaikavian is low at 20%. This is a hard Kaikavian dialect.

Serbo-Croatian has low intelligibility of Belorussian at 17%. Written is quite a bit better at 45% (Petričević 2019).

Serbo-Croatians understand Dubravica Kaikavian poorly at 17% (Petričević 2019). This is another hard Kaikavian dialect.

Serbo-Croatian comprehension of Prekmurjian is remarkably low at 15% (Petričević 2019). Only the Kaikavians seem to be able to understand this language.

Serbo-Croatians understand the hard Medjimurje Kaikavian dialect poorly at 15%.

Serbo-Croatian has only 15% intelligibility of Rusyn. Written comprehension is higher at 30% (Petričević 2019).

Serbo-Croatian understanding of the very archaic West Gorski Kotar Kaikavian is 13% (Petričević 2019).

Serbo-Croatian intelligibility of spoken Polish is poor, 12%* (Golubovic and Gooskens 2015a). Written is also very low, 16%* (Golubovic and Gooskens 2015a).

Serbo-Croatian understanding of Molise Croatian is extremely poor at 12%.

Serbo-Croatians barely understand Upper Sorbian at all at 10%.

Written is equally poor at 9%.

Serbo-Croatian understanding of written Lower Sorbian is the same at 9%.

Serbo-Croatian has [low intelligibility](#) of Russian at 10%. Oral intelligibility can go up to as much as 80% with bilingual learning. Written intelligibility is a bit higher at 17% (Petričević 2019).

Serbo-Croatian also has 8% intelligibility of Kashubian. Written Kashubian is not much better at 12% (Petričević 2019).

Serbo-Croatian has a very low 4% intelligibility of Rhodope Mountains Pomak.

Serbo-Croatian understanding of hard Zagorje Kaikavian is among the worst of all at 3% (Petričević 2019).

Serbs understand the hard Podravina Kaikavian just as poorly at 3% (Petričević 2019).

Bednjanski Kaikavian is unintelligible to most Štokavian speakers, with only 1% intelligibility (Petričević 2019).

Serbo-Croatians can't make heads of tails of Resian at 1% (Petričević 2019).

Serbo-Croatian understanding of Archaic Islander Čakavian is zero (Petričević 2019).

Serbo-Croatians understand nothing of Dolinci Burgenland Croatian (M. Jembrih 2019).

Bosnian

Bosnian is not actually a language at all. Instead it is a dialect of the Serbo-Croatian language called Neo-Štokavian. But because it is standardized, codified, has its own literature, and is the official language of a nation, it is worthwhile to discuss it in a separate section.

Bosnian is spoken throughout Bosnia and Herzegovina. It has two forms, Old Štokavian and New Štokavian.

Bosnian Old Štokavian has two forms, Bosnian Slavonian Old

Štokavian and East Bosnian Old Štokavian.

Derventa Slavonian Old Štokavian is a subdialect of the Slavonian Old Štokavian dialect spoken mostly in the Slavonian region of Croatia. It is spoken in Derventa, an enclave in the Republika Srpska in Bosnia. This area was formerly 40% Croatian, 40% Serbian, and 10% Bosnian, but due to ethnic cleaning during the war, it is now 80% Serbian, 10% Croatian and 10% Bosnian. There is still considerable discrimination and even some violence against the remaining Croats and Serbs by the ultranationalist Serbs who live here.

Derventa was probably brought here by Croats who came from Slavonia, and then it probably disseminated to the Serbs and Bosnians. It is now spoken here by all ethnic groups.

Jekavian-Šćakavian Old Štokavian or East Bosnian Old Štokavian is the most commonly spoken dialect in Bosnia. It is apparently fully intelligible with the rest of Bosnian. This language is spoken in a compact area in Central Bosnia around Sarajevo, Zenica, and Tuzla, where it is spoken by Bosniaks or Bosnian Muslims and by Croats and Serbs in Vareš, Usora, and other places.

It is interesting because it has *je* as its most common *yat* reflex. In addition to the basic Jekavian pronunciation, it also has mixed Ekavian–Jekavian pronunciations in Tešanj and Maglaj. For example, the following word can be pronounced two ways: *dete–djeteta*. The first form is Ekavian and the second form is Ijekavian. In Žepče and Jablanica, mixed Jekavian-Ikavian forms of the same word are possible: *djete–diteta*. Here the first form is Ekavian, and the second form is Ikavian. In the central area of this speech, the high back vowel *u* has changed to the diphthong *uo*. For example, *u > uo: vuk, stup > vuok, stuop*.

Bosnian New Štokavian has two forms, Bosnian Ikavian Štokavian and Bosnian Ijekavian Štokavian.

Bosnian Ijekavian Štokavian is a dialect of Serbo-Croatian and the basis for the standard language. It has full intelligibility with the rest of Serbo-Croatian. Although Croats' intelligibility of Bosnian is full, it is a bit lower than their intelligibility of Serbian, and some Croats have marginal to near-marginal understanding of Bosnian. A Croatian informant who had marginal intelligibility of Bosnian told me that this was due to the many Persian, Arabic, and Turkish loans in Bosnian

with which he was not familiar. This dialect is spoken throughout Bosnia and Herzegovina.

Bosnian Ikavian Štokavian is spoken in Western Bosnia east to Zenica, in the northwest around Bihać; in scattered parts of Northern Bosnia including the city of Banja Luka; and in Herzegovina in the Dalmatian Highlands. In Herzegovina, it is still widely spoken by Croats. It was formerly spoken by both Bosniaks and Croats in Western Bosnia in Travnik, Jajce, Bugojno, Vitez, but is dying out rapidly here. It is still spoken a lot around Bihać by Bosniaks. This dialect is different, but it is perfectly intelligible to other Bosnians, therefore it is not a part of the Dalmatian Ikavian Štokavian language.

Bosnian intelligibility of other Slavic lects:

A dearth of Bosnian informants meant few intelligibility figures for this language.

Bosnian has 100% intelligibility of Serbian.

Bosnian has 5% intelligibility of Medjimurje Kaikavian. Bosnians understand Medjimurje quite a bit worse than Croats do.

Montenegrin

Montenegrin, like Bosnian, is not an actual language. It is actually a dialect of the Serbo-Croatian Štokavian language. But it is the official language of a state, it is codified and standardized, and it has its own literature, so it is worthwhile for it to have a section of its own.

Montenegrin is spoken throughout Montenegro in two forms: Old Štokavian and New Štokavian.

Montenegrin New Štokavian has one form in Montenegro, Montenegrin Ijekavian Štokavian.

Montenegrin Ijekavian New Štokavian is the basis for *Standard Montenegrin*. It is based on East Herzegovinan Ijekavian New Štokavian. All other Serbo-Croatian groups have full intelligibility of this speech, but Croats have lower intelligibility of it than Serbs. Some Croats have marginal to near-marginal intelligibility of Montenegrin. Croats have slightly more problems with Montenegrin than with Bosnian.

Croats' problems with Montenegrin are the same as with Bosnian and are related to the huge number of Turkish, Persian, and Arabic loans in the language, the number of which is even greater than in Bosnia. There is a large Muslim population here whose speech was the source of these loans.

Croatia presently has few if any Muslims living there. Croatian Catholics tended to flee the Turkish invasions, while it appears the Orthodox Serbs and Montenegrins were either convert or stay and refuse to convert. Either way they would get heavily exposed to the Muslim speech with all of the Eastern loans. The Orthodox in Montenegro and Croatia seem to have converted more than those in Serbia, and Croats either fled or stayed and hardly converted at all.

Serbs have almost complete understanding of this dialect for some unknown reason, but they also have complete understanding of Bosnian. Apparently Serbs find Turkish, Persian, and Arabic loans easier to understand than Croats do. There is a long-term resident population of Serbian Muslims living in the Sanjak in Southwestern Serbia.

Montenegrin Old Štokavian has a single form in Montenegro, Zeta–Raška Old Štokavian.

Zeta–Raška Old Štokavian or Đekavian-Ijekavian Old Štokavian is one of the most diverse forms of Štokavian spoken in the Serbo-Croatian speech area, second only to Ikavian Štokavian. This dialect is spoken in the southeastern half of the Montenegro from the coast high up into the mountains to the border with Serbia. Most speakers in Montenegro are Muslims.

I am not certain why this language is called Đekavian-Ijekavian. Yes, there is the strange Đekavian reflex of *yat* in this dialect, but there is no Ijekavian reflex to be seen. Instead, there are Jekavian, Ekavian, and Ikavian *yat* reflexes. A better name for this dialect may be Đekavian-Jekavian, taking note of the odd Đekavian reflex while also noting that Jekavian is the most common reflex.

The most common *yat* reflex is Jekavian, next is Ekavian mixed with Đekavian, followed by Ikavian mixed with Đekavian. But the most common Jekavian reflex is not present in the name of the language.

In addition to the typical Jekavian speech, there are also mixed forms like Jekavian–Ekavian around Bijelo Polje: *djete–deteta*. Here the first

word is Jekavian, and the second word is Ekavian. Around Podgorica, there is an Ikavian-Đekavian mixed speech. Here the same word is *dite-đeteta*. The first word is Ikavian and the second word is the odd Đekavian. In the village of Mrkojevići in Southern Montenegro, Ekavian-Đekavian mixed forms are spoken: *dete-đeteta*. Here the first word is Ekavian, and the second word is Đekavian.

In some dialects, the yat reflex is a very open ϵ or \ae : *san, dan* > *sæn, dæn*. If these were more common, perhaps we might call them *ekavian* or *ækavian*. These reflexes are only very rarely found in the rest of Štokavian. In addition, *a + o* > \bar{a} or *a:* *kao, rekao* > *kā, rekā*. Here the *a + o* sequence went to long *a*.

Standard Croatian

Standard Croatian is part of the same Serbo-Croatian language that is also spoken in Serbia, Montenegro, and Bosnia/Herzegovina. This is all one language and not separate languages as politicians insist. These dialects were divided into separate languages based on politics instead of science.

However, Croatian can also be seen as a macrolanguage consisting of at least an incredible 12 different languages: Standard Croatian, Dalmatian, Bunjevac, Čakavian, Gan-Veyān Archaic Islander Čakavian, Rab Island Archaic Islander Čakavian, Kyrška Bešeda Archaic Islander Čakavian, Sušak Island Archaic Islander Čakavian, Komyzjanski Archaic Islander Čakavian, Korzulot Archaic Islander Čakavian, Molise Croatian, and Burgenland Croatian.

The following will deconstruct the popular notion that there is a language called Croatian, with dialects including Kaikavian, Čakavian, Štokavian, Molise Croatian, and Burgenland Croatian. The focus will be on the first three – Štokavian, Čakavian, and Kaikavian and will show that these are three separate languages, not dialects of a single tongue called Croatian. In fact, there is no language called Croatian, and the lect referred to as Standard Croatian is a dialect of a language called Štokavian. In this paper however, I moved Kaikavian out of Macro-Serbo-Croatian altogether and into Macro-Slovene.

Prominent Serbo-Croatian linguists now state that there are three

Croatian languages: Croatian, Kaikavian, and Čakavian (Kapović 2010, Silić 1998). However, judgments by Croatian linguists must be taken with a grain of salt, as linguistics was badly corrupted by politics during the Communist era, and many Croatian linguists continue to be intellectually corrupted, at the moment by Croatian nationalism. Like many other Slavic nations, nationalism has badly damaged linguistics in this country. In addition, Croatian nationalists are some of the worst nationalists in the Slavic world, and they have a lot of competition.

These three main Croatian dialects are all completely unintelligible with each other. Some Croatian nationalists insist that all of these are “dialects of Croatian” and that “Croatian” is a single language with these three dialects. They also insist that MI between all of these dialects is full and that most Croatians can easily comprehend all of these dialects. I know firsthand that this is not true because I have had informants who understood 0-3% of both Kaikavian and Čakavian. These same informants also told me that most Croatian Štokavian speakers simply do not understand these dialects at all.

For one, the lexicons of Kaikavian and Čakavian differ dramatically from Štokavian (M. Jembrih 2019). A comparison of a few words in Kaikavian and Štokavian below makes this very clear:

“house”

Kajkavian: *hiža*

Štokavian: *kuća*

“to work”

Kaikavian: *delati*

Štokavian: *raditi*

“to run”

Kaikavian: *bežati*

Štokavian: *trčati*

“stairs”

Kaikavian: *štenge* (German borrowing *Stiege*)

Štokavian: *stepenice* (M. Jembrih 2019).

One of the major differences between Štokavian and Kaikavian is in the lexicon. Many Kaikavian words are not in Štokavian. Instead, these words resemble Slovene. However, Kaikavian also has words that are in neither Slovene nor Kaikavian (M. Jembrih 2019).

On the other hand, all of the different kinds of NeoŠtokavian – Ekavian, Ikavian, Ijekavian, and Jekavian differ in their yat reflexes, but in general the vocabulary is more or less the same in the various Neo-Štokavian dialects. You can see above how Kaikavian dramatically differs from the dialects of Štokavian. How then can Kaikavian be a dialect of Štokavian or Serbo-Croatian? It makes no sense at all (M. Jembrih 2019).

Now that we have settled the question of the fake language called Croatian and its three “dialects,” let us now look at the actual linguistic landscape in Croatia.

10% of all Croats speak Ikavian Štokavian as a mother tongue. It is spoken in Dalmatia, along the coast from Dalmatia north to the Istria Peninsula in the former Čakavian speaking area, and in the Lika Highlands. It is being replaced by Ijekavian Štokavian in urban areas, but it is increasing in the south by supplanting Čakavian.

56% of Croats speak Ijekavian Štokavian as a native language, which is close to the Serbo-Croatian Standard. It is increasing due to schooling and the media. Mostly they speak native Croatian forms of Ijekavian Štokavian often referred to as Western Ijekavian Štokavian.

Fully 42% of Croats either do not speak Serbo-Croatian or do not speak it well. Many of them do not even understand Serbo-Croatian well. Instead, they speak Čakavian, Kaikavian and Torlakian.

31% of Northern Croats speak Kaikavian, mostly in the north. Kaikavian is losing ground to Štokavian in the southeast around Slavonia but it is gaining ground against Čakavian in the southwest.

11% speak Čakavian along the Adriatic coast and on the islands, on the Istrian Peninsula, in Dalmatia, inland at the Gacka Valley, and in some other places. Čakavian is definitely losing speakers.

2% of Croatians speak Torlakian, mostly in Zagreb and Northern Dalmatia. These are Catholic Serbo-Croatians who were living in Kosovo and Romania who immigrated to Croatia and thereby became Croatians.

The new standard Serbo-Croatian was chosen in 1850 on rather hazy grounds.

It has little to do with the languages traditionally spoken in Croatia. The true language of Croatia is probably something like a Old Štokavian-Kaikavian-Čakavian mix, all of which were very similar for centuries before the divisions of the region under new rulers (Petrov 2015).

Croatian nationalists support Štokavian as the national language of Croatia on flimsy grounds. For instance, they claim that Kaikavian and Čakavian were not chosen because they had too many foreign words (Kaikavian had many German and Hungarian loans, and Čakavian had many loans from Romance). However, this argument made no sense because Štokavian, which was preferred, was also full of loans, in this case Turkish and other vocabulary from the east (Petrov 2015).

Croats were also very unhappy with the Vuk's new Serbo-Croatian standard in which the two main languages of Croatia, Kaikavian and Čakavian, were negated in favor of Štokavian, a language that few Croats spoke. Before 1850, most peasants simply spoke their local Croatian dialect of Kaikavian or Čakavian.

However, educated Croats typically spoke Latin, German, whatever the local Croatian dialect was, and to a lesser extent Hungarian. Much writing was in German or even Latin, but there was also a long tradition of writing in local Croatian, particularly in Kaikavian, which was more or less the official Croatian literary language before 1850.

Before 1850, the language of the Croatian Parliament was, incredibly enough, Latin. Most books for sale in Zagreb were in German (M. Jembrih 2019). All of these languages continue to affect local Croatian languages, as Čakavian is full of Latinisms, and Kaikavian is full of German and Hungarian loans, particularly in the main Kaikavian region in the east.

So after 1850, the local Slavic language spoken by nearly all Croatians was simply wiped out officially and replaced with the new Štokavian standard. This process began a long and serious decline in

Čakavian and Kaikavian.

Nevertheless, the new Štokavian Croatian standard is not spoken much outside of official functions. So once again we are confronted with the linguistic cliché of “an official language that no one speaks.”

In Croatia today, there are three types of Croatian spoken.

First, the official standard, which exists more on paper than in reality. In fact, if you speak only Standard Croatian, you will be regarded with puzzlement in most parts of Croatia where you will be probably seen more as an uneducated peasant from the hinterlands than as a modern educated person.

Second, a regional standard, nevertheless intelligible across the land, that is a mixture of Standard Croatian and aspects of local Kaikavian and Čakavian dialects. This second form is surprisingly common, even on television.

Third, the more pure forms of Čakavian and Kaikavian which tend to be spoken more among old people and in rural areas. Speakers are often ridiculed as uneducated peasants, and there is a lot of pressure to give up the hard traditional Kaikavian and Čakavian dialects.

The gap between Standard Croatian and the local mixed Štokavian forms continues to grow. There are many local words, often German, Turkish, or Italian borrowings, that would be acceptable in an informal context but would nevertheless be very inappropriate in a formal setting. By the same token, speaking Standard Croatian in an informal setting would be seen as stilted and odd.

In all of the former Yugoslavia, the gap between formal and informal speech is probably greater in Croatia than anywhere else.

Štokavian has many Turkisms and other Orientalisms in its vocabulary, while Čakavian has archaic Slavisms or Romance vocabulary, and Kaikavian and Slovene have Germanisms.

Most Croatian dialects have both pitch accent and long vowels which are complex, as they change within words with inflection. Croatians often refer to their pitch accent as “tones,” although formally the language has pitch accent and not actual tones.

Although vowel length is phonemic in Croatian, there are only a few minimal pairs that differ in vowel length, so in general it is not necessary to use either pitch accent or long vowels to be understood

when speaking Croatian, as speech that lacks these two things is still intelligible and only marks you as a 2nd language speaker.

As in the rest of Serbo-Croatian, the aorist and imperfect tenses are heading out in Croatian. The aorist is probably further gone in Standard Croatian than in any other type of Serbo-Croatian.

Štokavian has two types spoken in Croatia, Old Štokavian and New Štokavian.

Almost all Croatians speak New Štokavian, as Old Štokavian is only spoken in a few places in the far east near the Serbian border.

Croatian Old Štokavian has three dialects, Archaic Šćakavian Old Štokavian, Zeta-Lovcen Štokavian, and Torlakian Old Štokavian.

Archaic Šćakavian, Slavonian Štokavian, Osijek Štokavian, Osijek Croatian, Continental Štokavian, or Slavonian is an Old Štokavian Ekavian dialect spoken in the the Slavonian region in the east of Croatia. However, in Osijek, this dialect is spoken by only 25% of the population, with the other 75% being Ijekavian speakers. This is true "hard" Štokavian. This dialect is said to be [somewhat similar](#) to Kaikavian. Slavonian is both Ekavian and Ikavian. Ikavian predominates in Posavina, Baranja, and Vukovar, Ekavian is dominant in Podravina, and the speech in Požega is a mixed Ikavian-Ekavian dialect.

Posavian Old Štokavian or Posavian Slavonian is a Slavonian dialect thought to be the remains of the original Štokavian. It is now present only in Southeastern Croatia on the Bosnian border. It exists only in some villages and not to the extent it was before, but there are still Ikavian accents and Old Štokavian words. It is however still spoken in Vinkovci in Southeastern Croatia. The speech of the older people is Ikavian, while the younger people speak more Ekavian.

Pleternica Slavonian is an interesting dialect that has been studied by educational researchers studying how well children learn Standard Croatian. In Pleternica, students show up for school as Slavonian monolinguals and learn Standard Croatian as a foreign language. Pleternica is part of the Požega Slavonian dialect. The speech here seems to be Ikavian-Ekavian, but it was hard to determine this with accuracy as data was lacking. Whether this means that Posavian Slavonian part of a is a separate language or not is not known (Pavličević-Franić and Aladrović 2017).

Zeta Lovcen Old Štokavian is spoken in only a single town in Croatia, Peroj in West Istria. This is an ancient Montenegrin dialect brought by Montenegrin immigrants who arrived in the 1750's. It retains much of its ancient character. Užice-Zlatibor is also part of this dialect. The formal name for the entire dialect is Zeta–Raška Štokavian. This is one of the most different of the Štokavian dialects. It is probably best classed as Đekavian-Jekavian, which means that it is Ikavian-Jekavian, Ekavian-Jekavian and Jekavian-Ekavian.

New Štokavian Croatian is spoken in three forms in Croatia, two forms of Ijekavian Štokavian, a native Ijekavian form of various types and Standard Croatian Ijekavian, and one form of Ikavian Štokavian, Ikavian Štokavian or Younger Ikavian.

Croatian Ikavian New Štokavian has one dialect in Croatia, Western Ikavian New Štokavian.

Ikavian Štokavian Croatian, Younger Ikavian Croatian, or Western Ikavian Štokavian Croatian is a separate language, or at least the type spoken in Dalmatia and by the Šokci in Posavina in Croatia is. Šokački is definitely a separate language, part of the same language as Bunjevci. Dalmatian is a different language than Bunjevci/Šokački and is dealt with in the same section. It is not known if the Ikavian Štokavian spoken in the Lika Highlands and along the coast in the old Čakavian zone is part of this language or together with the Ikavian Štokavian in Bosnia as dialects of Serbo-Croatian spoken.

Croatian Ijekavian NeoŠtokavian is spoken throughout the country in both the official Standard Croatian form and a variety of native forms, the most well known of which is Dubrovnik Ijekavian Štokavian, a dialect of the Dalmatian language.

In the islands, native Ijekavian NeoŠtokavian is spoken in spots on Pag, Hvar, Brač, Korčula, and Šolta (Kalsbeek 1998).

Zagreb Štokavian, Zagreb Kaikavian, Koine Kaikavian, or Purgeri is a cultivated urban Kaikavized Ijekavian Štokavian dialect that is interesting in that it lacks pitch accent and vowel length (Smiljanic 2013). *Purgeri* is actually a German borrowing from the German word *Burger*, meaning citizen, which meant the residents of the city as opposed to rural people, so *Purgeri* simply means the people who live in Zagreb.

60% of the population of Zagreb, the capital of Croatia, or 410,000 people, say they speak Kaikavian at least at home with their families. This group is mostly in the center and northwest of the city. In the streets, tramways, and buses of the capital, the half-Kaikavian purgeri dialect is heard more often than Štokavian. All of the main suburbs of Zagreb - Zaprešič, Sesvete, Velka Gorica – are Purgeri-speaking, with 50,000 speakers in both Sesvete and Velka Gorica. Varaždin has 60,000 Purgeri speakers, including thousands of Agramer speakers.

This dialect, at least as spoken by the youth of the city, is modernized and under the influence of Štokavian. This dialect is now referred to as Koine Kaikavian, similar to the way an Ancient Greek variety is referred to as Koine Greek. The term *koinē* generally means a language, part of a larger language or macrolanguage, that is central enough that it can be understood by speakers of all lesser lects included in the the larger language or macrolanguage.

The best way to see this dialect is as Kaikavized Štokavian dialect, as it is now more Štokavian than Kaikavian, so it should actually be called Zagreb Štokavian instead of Zagreb Kaikavian. This is especially true in light of the fact that this Zagreb dialect is fully intelligible with Serbo-Croatian. However, it was originally a Kaikavian dialect that got so Štokavized that it turned into a Štokavian dialect.

This dialect is extremely popular. In fact, the young people relate so heavily to this dialect that there are gangs of Purgeri speakers who beat up Standard Croatian speakers in the capital. In the near future, it is possible that there will be two mutually intelligible standard languages in Croatia like Nynorsk and Bokmal in Norway – Standard Croatian and Kaikavian.

Despite the full intelligibility with Standard Croatian, Purgeri still exhibits [very large differences](#) with Standard Štokavian in grammar, lexicon, phonology, and accentuation.

Purgeri children show up in school monolingual in Purgeri and learn Standard Croatian more or less as a separate language (Pavličević-Franić and Aladrović 2017).

People are now migrating to Zagreb from all over Croatia and bringing their local dialects or languages with them. As a result the language of the city is becoming very mixed. All of these lects are mixing together and combining into Purgeri.

Purgeri is actually the prestige dialect of Zagreb. This is different from other large cities. For instance, recent immigrants to New York do not typically adopt New York English because it is not a prestige dialect. However, most new immigrants to Zagreb soon start speaking Purgeri because of its prestigious nature.

Although there are some questions about whether to place Purgeri in Štokavian or Kaikavian, and it was definitely a Kaikavian dialect in the past, its mutual intelligibility with Standard Croatian means it must be put in Štokavian.

Standard Croatian intelligibility of various lects:

Croatian intelligibility investigation was state of the art, with many Croatian informants and a couple of superb personal informants, including one linguist.

Croatian has 99% of Purgeri Štokavian (Petričević 2019).

Croatian has 97% of Standard Burgenland Croatian (M. Jembrih 2015). The standard Burgenland language does have some Old Kaikavian words that are not understandable to Croatian speakers except possibly in context (M. Jembrih 2015). The various Burgenland Croatian dialects, however, have poor intelligibility with Croatian. Burgenland Croatian speakers feel that their dialect is different from Croatian for various reasons, including sociolinguistic factors.

Croatian intelligibility of Serbian is 95% (Petričević 2019).

Although Croatian intelligibility of Bosnian and Montenegrin is full, Croatians have a bit of a harder time understanding it than they do with Serbian due to the many Arabic, Persian, and Turkish loans in the languages, with which they are not familiar.

Croatian intelligibility of Bosnian is 93% (Petričević 2019).

Croatian intelligibility of Montenegrin is 92% (Petričević 2019).

Croatian intelligibility of Virje-Djurdjevac Kaikavian is 88%. This is a Kaikavian dialect transitional to Serbo-Croatian.

Croatian intelligibility of Šokački is 81%. However, written understanding is full at 93% (Petričević 2019).

Croatian intelligibility of Ugljan Island Čakavian is quite high at 75% (Petričević 2019).

Croatian has 60% intelligibility of Torlakian (Petričević 2019).

Croatian intelligibility of Slovene is good at 55% (M. Jembrih 2019). Croatian has 46% understanding of Virovitica-Podravina Kaikavian. Croatian an intelligibility of 39% of Kaikavian, but the range is very high (range 3-88%). Written is about the same at 40% (Petričević 2019).

Croatian intelligibility of Čakavian is 37% (M. Jembrih 2015), but the range is high (0-75%).

Croatian intelligibility of Macedonian is 25% (Petričević 2019).

Croatian understanding of Medjimurje Kaikavian is 25%.

Croatian has 20% intelligibility of Belarussian. Written is higher at 45% (Petričević 2019).

Croatian has 20% intelligibility of Ukrainian. Written is significantly better at 35%.

Croatian has 17% intelligibility of Dubravica Kaikavian (Petričević 2019).

Croatian intelligibility of Rusyn is poor at 15%. Written is a bit better at 30% (Petričević 2019).

Croatian understands Prekmurjian at only a rate 15% (Petričević 2019).

Croatian intelligibility of Molise Croatian is very poor at 12%.

Croatian has poor intelligibility of Russian at 10%. Written is also quite poor at only 17% (Petričević 2019).

Croatian has remarkably poor intelligibility of Kashubian at 7%. Written intelligibility is also very poor at 15% (Petričević 2019).

Croatian intelligibility of Zagorje Kaikavian is remarkably low at 3% (Petričević 2019).

Croatian intelligibility of Bednjanski Kaikavian is extremely low at 3% (Petričević 2019).

Croatians cannot understand Archaic Islander Čakavian at all (Petričević 2019).

The Dolinci Burgenland Croatian Old Čakavian dialect spoken in Unterpullendorf cannot be understood at all by Croatian speakers (M. Jembrih 2015).

Murláška Besýda was spoken by a small shepherd community on Bleynatórna Mountain on the border of Southwestern Bosnia and the Dalmatian mainland in Croatia. It was an archaic mix of an Ikavian dialect of Old Štokavian and Romanian and was not intelligible with any other Slavic or Romance lect. It was about halfway between Serbo-Croatian and Romanian. Although it had 1,200 speakers before WW1, this language has now gone extinct.

Croatian Western Ikavian

Croatian Western Ikavian, Croatian Ikavian Štokavian, Young Ikavian Štokavian, or Dalmatian-Bosnian Štokavian is a group of Štokavian dialects spoken in various forms mostly in Croatia, Western Bosnia, and part of Vojvodina in Serbia, at least some of which seem to form two actual separate Štokavian languages.

At least the Ikavian Štokavian lects spoken as Dalmatian Croatian by Croats in Dalmatia and as Šokački by the Šokci Croats in Posavia in Southeastern Croatia and Vojvodina in Serbia, and as Bunjevaci also in Vojvodina appear to be separate languages.

Other Croats have a very hard time understanding hard Dalmatian dialect, and Croats and Serbs have some difficulty understanding Šokački and Bunjevaci respectively, so it makes sense to split off at least this part of Ikavian Štokavian off as a separate language. But Dalmatians cannot understand Šokački themselves, so it appears that we are dealing with two separate languages here, Dalmatian and Bunjevaci, with Šokački being a part of Bunjevaci.

Ikavian Štokavian is spoken as a mother tongue by 10% of Croats in a wide band along the border with Bosnia and Serbia and inland for quite some ways to the north on the southern end and to the west on the eastern end. It is spoken in the cities of Dubrovnik, Cavtat, Osijek, and Slunj. Most of this speech is part of Standard Croatian, but the part spoken in Dalmatia is part of a separate language called Dalmatian and the type spoken by the Šokci is part of another language called Bunjevaci.

The Ikavian Štokavian spoken in the Lika Highlands of the Gacka Valley lacks intelligibility data, so we don't know if it is part of this

language or not.

Dalmatian

Dalmatian or Western Ikavian Štokavian is a separate Ikavian Štokavian spoken in Croatia in Dalmatia and the surrounding area. It is a Štokavian language transitional to Čakavian.

Zadar Dalmatian is a Western Dalmatian dialect spoken near the Čakavian-speaking area in Zadar near the Dalmatian coast. It is an example of a Štokavian dialect that has undergone heavy Čakavian influence. It is much closer to Čakavian than the more inland or "Continental Štokavian." It is particularly close to Uglyan Island Čakavian. Zadar Štokavian still has Čakavian elements and also retains many Romanisms. Fifty years ago, all coastal cities were Čakavian speaking, but now some, such as Zadar, have gone over to Štokavian (Kolanović 2015).

After World War 2, many Italians left or possibly fled the area as Communist Partisans purged Italians or people thought to speak an Italian-type Čakavian dialect such as Archaic Islander Čakavian. They were replaced by Štokavian speakers from the rural Croatian hinterlands, hence Zadar shifted from Čakavian to Štokavian (M. Jembrih 2019). The population here is 50% Dalmatian, with most of the remainder being former Čakavian speakers from the island now speaking Ikavian Šćakavian, along with many Ijekavian Štokavians.

At least Slavonian Štokavian speakers have a hard time understanding Zadar.

Šibenik Dalmatian is another Dalmatian dialect transitional to Čakavian spoken in the city of Šibenik near Dalmatia. Šibenik stopped speaking Čakavian long ago, around 1500's-1600's. Šibenik is quite similar to Zadar (M. Jembrih 2019). The speakers in the city are Dalmatian and Ikavian-Šćakavian, with the mix mirroring that of Zadar. We lack intelligibility figures for Šibenik, but it seems to be part of this language.

Pula Dalmatian, spoken in the city of Pula in the southeast of Istria, has the same situation as Zadar, Split, and Šibenik, with most of the

population being Dalmatian and the remainder being the usual Ijekavian Štokavian and Ikavian Šćakavian mix. Although we lack intelligibility data on Pula, it seems to be part of the Dalmatian language.

There was also a Romance language, Pula Venetian, a dialect of Venetian similar to Fiume and Triestine, spoken here, which is either extinct or there are only a few speakers left. This Adriatic Venetian is intelligible across the region from Pula to Fiume. There are also some speakers of Pula Venetian in the diaspora, including middle-aged speakers in the US.

Split Dalmatian spoken in the Dalmatian city of Split, is similar to Šibenik and Zadar, being an Dalmatian dialect with considerable Čakavian influence. Most speakers here, 75% or 150,000, are Dalmatians by mother tongue, this being their largest conglomeration in Croatia. Most of these speakers came from the Dalmatian highlands. The rest of the population is Ijekavian and other groups. Only 100 years ago, Split was majority Čakavian. The Čakavian immigrants from the nearby islands have mostly become Ikavian-Šćakavian speakers.

Oddly enough, this language also has an Ijekavian variant that appears to be part of the same language as the Ikavian variant.

Dubrovnik Dalmatian is a Western Ijekavian Štokavian dialect of Dalmatian. The morphology is Ijekavian Štokavian, but it is full of Romance vocabulary due to Čakavian influence (Kolanović 2015). The Romance loans are from the Ragusa dialect of the extinct Dalmatian language and from Italian and Venetian. It is 75% Ijekavian Štokavian, with the remainder being Čakavian Islanders now speaking an Ekavian Šćakavian dialect. In the Middle Ages, speakers here spoke the exotic Ijekavian Čakavian, now limited only to Lastovo island and Janjina on the Pelješac Peninsula.

It is spoken all along the Dubrovnik portion of Croatia south of the Bosnian border from Štedrika south to Molunat, on Kolučap, Lopud, and Otok Šipán Islands north of Dubrovnik and on the eastern half of the Pelješac Peninsula. Just to the west, in the center of the peninsula, a unclassified mixed Southern Čakavian-Dubrovnik Štokavian dialect is spoken. The western half of the peninsula speaks Southern Čakavian.

Cavtat Dalmatian is a Western Ijekavian Štokavian dialect of Dubrovnik Dalmatian spoken in the city of Cavtat a few miles from Dubrovnik. 10 years ago, this was one of the last holdouts of very pure Dubrovnik dialect (M. Jembrih 2019), and the dialect seems to be still going very strong there. This dialect continues to be spoken by all ages by most everyone in the city. Ijekavian Štokavian speakers from Slavonia often [cannot understand](#) the Ijekavian Štokavian speakers of Cavtat. Even after eight years living there, they still have a [hard time](#) understanding the locals at times.

The dialect is hard to understand due to numerous archaisms and a very large number of Romance loans (M. Jembrih 2019).

[*Watsonville Konavle Dalmatian*](#) is a dialect of Dalmatian spoken in Watsonville, California (Filipovic 1996), consisting mostly of Cavtat Dalmatian along with much smaller amounts of two nearby similar Croatian languages.

It has been that has been spoken in the US for three generations (Filipovic 1996). They came to the US from the Konavle region, part of Dubrovnik-Neretva County just southeast of Dubrovnik, which includes the main cities and towns of Cavtat, Čilipi, Molunat, Pločice, Popovići, Močići, Gruda, and Ljuta along with many more small villages. The northern border of the region begins eight miles southeast of Dubrovnik.

Others came from the island of Brač. Both Brač Southern Čakavian, a dialect of Southern Čakavian and Brač Archaic Islander Čakavian, a dialect of Archaic Islander Čakavian, are spoken there, but we do not know which of these languages the islanders spoke. Archaic Islander Čakavian, Southern Čakavian, and Dubrovnik Dalmatian are not mutually intelligible. However, it appears that the vast majority of the migrants came were Dalmatian speakers from Konavle, with the speakers of the Brač Island languages being a much smaller group.

A web page [devoted to the language](#) listed many members of this community down through the years, and all were from Konavle. All three of these languages may have combined to form some sort of a koine in Watsonville, but this is not proven. It is only the best theory. All three languages do have something in common. The Čakavian languages are part of Macro-Čakavian, even if they are hardly intelligible. And Dalmatian has a huge Čakavian substrate. So these language did have things in common and could conceivably formed

some new odd koine in the US.

Speakers came from the larger Konavle cities and towns of Čilipi, Popovići, Gruda, and Ljuta, while others came from smaller Konavle villages of Uskoplje, Lovorno, Pridvorje, Dunave, Zastolje, Bani, Komaji, Drvenik, Vodovod'a, and Mihanići. Quite a few others came from other parts of Dalmatia, including Klis, and Mijkovi, Gromača, and Slano in the Dubrovačka Primorje region. A few others came from Škrip on Brač Island.

The first migrants arrived in the 1880's, but the largest group came between 1900-1914.

In 1992, three generations had spoken Konavle and the dialect was still being maintained well (Filipovic 1996).

These people call themselves *Konavljani*. Most of the Konavljani settled in the Pajaro Valley, a rich farming area. Almost all of the Konavljani ended up being farmers. Others worked in area factories. The second and third generation Konavljani also worked in the fields on their families' farms. The parents' generation of the first generation mostly never learned to speak English. All communication on the farm were in Konavle. Konavljani workers in the factories never learned to speak English either. The other workers in the factories also spoke Konavle, and communication with management was maintained by a Konavljani who was a Konavle-English bilingual (Filipovic 1996).

The second generation children of the first generation Konavljani showed up at school as Konavle monolinguals. A number of these people were still alive and working in professions when interviewed in 1992. They were true bilinguals in English and Konavle. Their Konavle was excellent, with pure pronunciation and rich vocabularies (Filipovic 1996).

Konavle has borrowed quite a few words from English such as *car, boy, shoe, trouble, build, like, teach, and walk*. It is not known why they borrowed these English words as clearly Konavle had words for all of them except possibly *car*. Some of these speakers sometimes borrow a word in English as it has a much richer vocabulary than Konavle. The second and third generation Konavljani learned English, but they continued to work in the fields which were very conducive to maintenance of Konavle as most of the workers were Konavljani who also spoke the dialect, so they all became full bilinguals (Filipovic

1996).

The main factor in the maintenance of the language was the family. Families are very tight-knit and parents have instilled in their children a love of Konavle that they continued to reinforce even as their children became young adults (Filipovic 1996).

There is no information on whether Watsonville Konavle Croatian is still intelligible with Konavle Dalmatian back in Croatia.

Bunjevaci

Bunjevaci, Šokački, or Bunjevaci/Šokački, Šokački Ikavian Štokavian, or Bunjevaci Ikavian Štokavian is a separate Ikavian Štokavian language spoken in Southeastern Croatia by a people called Šokci and in the Vojvodina region of Serbia by a people called Bunjevacs. It is a separate language from Dalmatian since Dalmatian speakers cannot fully understand it.

Šokački Ikavian Štokavian is spoken by Šokci Croats in a small part of Southeastern Croatia near the Bosnian and Serbian borders. The Šokci are a specific ethnic group with their own outfits, customs, music and dialect. The Šokački dialect is very similar to the Bunjevaci dialect spoken in Serbia. In fact, the Bunjevacs may have been Šokci Croats who migrated to the far north of Serbia long ago.

Bunjevaci Ikavian Štokavian is spoken by Croat migrants in Serbia in a few places in Vojvodina and across the border in Hungary.

The Bunjevacs were originally Catholic Croats and Bosnian Croats who fled east during the Byzantine wars waged to reconquer these areas from the Turks. They were part of another migration of Balkan Christians fleeing north and east during the Ottoman Empire, especially during wars and uprisings.

Probably 90% of Bunjevaci speakers identify as Serbs, and maybe 10% as Croats. Bunjev Serbs resent what they call attempts by the Croatian government to "Croatianize" them. Speakers say they speak a separate language and want it to be recognized. A dictionary of 20,000 Bunjevaci words has been compiled, and the grammar is somewhat different than Serbian Štokavian.

Serbo-Croatian linguists usually say that Bunjevac is a dialect of Serbo-Croatian.

Be that as it may, further investigation reveals that not only is Bunjevac the same language as spoken by the Šokci in Southeastern Croatia but it is also similar to the Ikavian Štokavian that is spoken around Dalmatia and in Western Bosnia. But [Serbs have a hard time](#) understanding Bunjevac, while Croats have a hard time understanding Dalmatian Šokački. Furthermore, Dalmatians have a hard time understanding Šokački.

Some Serbs say the Bunjevac "speak like aliens." Serbs definitely regard Bunjevac as one of the strangest and most difficult dialects in the country, perhaps second only to Torlakian, which many Serbs seem to regard as a completely separate language.

Bunjevacs say that if Bosnian, Croatian, Montenegrin, and Serbian can all be separate languages while being mutually intelligible, why can't Bunjevac be a language just like the other four languages. At first this seems silly as if every little dialect in Serbo-Croatian region is trying to codify their particular speech as a standardized full language. Obviously things could get very silly quite fast.

What is shocking though is that of all five languages, Bunjevac and the four standard Serbo-Croatian languages, Bunjevac has the best argument for being a standardized language of all of them because it is the only one that lacks full MI with all of the rest. In other words, the four standard languages are not really true languages and are better seen as four dialects of a single language that have codified themselves as separate languages due more to politics, ethnic nationalism, and brutal ethnic warfare than to linguistic science.

On the other hand, the Bunjevacs, who are ridiculed by the others as trying to standardize and codify what is nothing more than an inconsequential dialect of Serbian, are the only group with a linguistically justified claim to a separate language. The Bunjevac language surely deserves to be recognized as a separate language and codified on the grounds of linguistic science. The only reason not to do so is Serbian nationalist politics.

Čakavian

Čakavian is said to be a dialect of Serbo-Croatian, but this is incorrect and has more to do with politics than Linguistics, as it is a full language. In fact, it is better seen as a macrolanguage consisting of five separate languages – the main Čakavian language and at least four languages inside a second branch of Čakavian called Archaic Islander Čakavian. Čakavian has 500,000 speakers, mostly along the coast of Croatia and in the offshore islands. 11% of the population of Croatia speaks Čakavian. Čakavian has been in retreat for some time and it is continuing to lose speakers.

Čakavian is spoken in almost all of Istria; on all the islands of the Adriatic coast with the exception of the southeast around Mljet and surrounding small islands, a coastal strip known as Primorje Croatia; the town of Senj; narrow, non-contiguous strips along the Dalmatian coast around Zadar, Split, and Šibenik; scattered areas in inland Croatia, especially from Otočac northwards; parts of Austria, Hungary, and Slovakia (as Burgenland Croatia); and three villages in the province of Italy (as Molise Slavic) (Kalsbeek 1998).

Čakavian differs from the other nearby Slavic languages due to the presence of many Italian words. In fact, 25-50% of the vocabulary is from Romance. Čakavian is by far the most Romanized Slavic language.

Čakavian has been in retreat for a long time, centuries actually, for a wide variety of reasons. Presently it is only spoken by 7% of Croats. Čakavian appears to be slowly dying out. Many of the most fluent speakers are older, and Čakavian is becoming increasingly Štokavized. There are cities on the coast of Croatia such as Dubrovnik that were Čakavian-speaking only 50 years ago that now speak more or less a normative Serbo-Croatian. Nevertheless, there are still young adults who are fluent speakers.

Nevertheless, Čakavian speakers are quite busy trying to save their language. Dictionaries of many Čakavian dialects have already been made. On the Istrian Peninsula, there is now a Čakavian Parliament dedicated to saving the language. Istrian Čakavian in general is closer to Kaikavian than other Čakavian dialects are (Petrov 2015).

Čakavian magazines are published even today (M. Jembrih 2014).

Čakavian may have been much more widely spoken in the past. Renowned Bosnian linguist Asim Peco did a lot of research on Old

Čakavian-Ikavian Štokavian dialects in Western Bosnia. In his work he introduced the term *Šćakavian*, apparently for mixed Štokavian and Čakavian dialects. In his works, Peco suggests that Čakavian was once much more widely spoken along the Croatian coast, whereas nowadays it is spoken mostly on the islands.

Inland and to the east of this area was a large region in which Šćakavian or Čakavian-Štokavian transitional dialects were spoken. This was part of a dialect continuum between Čakavian and *Old Štokavian*. It is important to note that this continuum was with Old Štokavian, not with Novoštokavian. Old Štokavian was very different from Novoštokavian, the basis for modern Serbo-Croatian.

Nowadays, these dialects are still present, however they are now heavily Štokavized.

The only significant difference between Šćakavian and *Kvarner Archipelago Čakavian* is that *ča* for the interrogative pronoun “what” has been replaced with *što* (M. Jembrih 2014). Of course this classic replacement indicates Štokavian influence.

However, it should be noted that presence of either the *ča* or *što* pronouns in a Croatian lect does not alone tell us whether it is Čakavian or Štokavian (M. Jembrih 2014), although it would seem intuitive that it would.

The etymology of the *ča* or *što* pronouns is interesting. The change was probably *Šta* > *Šća* > *Ča*. At first the alveolar stop was palatalized under the influence of the preceding palatal. Then the two alveopalatals next to each other reduced to one. At any rate, *Ča* is certainly related etymologically to *Šta*, whereas *Kaj/Kai* is not (M. Jembrih 2014).

The linkage of Croatia and parts of Bosnia goes back quite a ways. Hrvoje Vukčić Hrvatinić was a ruler of Croatia, and during his rule, Croatia included the parts of Bosnia and Dalmatia where these Šćakavian dialects are now spoken. The Hval Manuscript which is now stored in Bologna, Italy, and is dedicated to Hrvoje Vukčić Hrvatinić is rumored to be written in this old Šćakavian dialect, but this has not been confirmed (M. Jembrih 2014).

Čakavian is closely linked with Croatia, whereas Kaikavian is more linked to Slovenia. Bosnia is where the Balkan Region formally begins, whereas areas to the north see themselves as more connected to

Central Europe (M. Jembrih 2014).

In addition to Peco's findings, there are interesting reports of military texts from the 1500's and 1600's during Turkish incursions to the area and battles around Una and Bihać in Western Bosnia that are said to be written in pure Čakavian. Although searches for these reports have come up empty-handed, there is a report that one of these reports may be presently housed in a military museum in Vienna.

Čakavian has a long history as a written language, but it was mostly written down long ago. From the 1100's-1700's, Čakavian was written in peculiar Glagolitic script. Although many existing old Croatian Glagolitic texts are written in OCS, in the main Čakavian area of Istria and Kvarner, only 20% of the texts are in OCS, and the rest are in Old Čakavian. One of the earliest Croatian texts of all - the Baška tablet from 1105 (Bujan 2015), is written in what looks like *Old Čakavian*. Elderly Gan-Veyan Čakavian speakers from Jurandvor, where the tablet was found, were able to recognize all of the words on the tablet except two.

From the 1100's-1500's, Čakavian was the principal and official spoken language in the Medieval Croatian Kingdom. At that time, 2/3 of all Croats spoke Čakavian. It was also widely spoken along the Montenegrin coast and in Western Bosnia. Writing in Čakavian began to slow down in the 1500's when writing in Kaikavian began to rise.

From the 1500's-1700's, the Čakavian area was invaded by Turks. Many Čakavian speakers fled to Burgenland in Austria or across the Adriatic to Italy, where they settled around Molise. Burgenland Croatian and Molise Slavic, spoken to this day, are left over from this migration. The remains of the Čakavian speakers in former Čakavian-speaking zone then mixed with Štokavian speakers from Sandjak, Bosnia and Kosovo who arrived with the Turks. The result was the Ikavish-Sćakavian uncovered by Peco and discussed above.

Due both to its exotic script and the fact that it was somewhat isolated from the body of Slavic, Čakavian has retained many archaic Slavicisms. For instance, Čakavian has many words that are the same as or similar to Polish, Ukrainian and Slovak words. In Serbo-Croatian, these terms are called archaisms and are substituted for with Turkisms.

Although Čakavian is clearly a separate language from Croatian, in

Croatia it is said that there is only one Croatian language, and that is Standard Croatian, Croatian Štokavian, or Serbo-Croatian. The idea is that the Kaikavian and Čakavian languages simply do not exist (M. Jembrih 2014).

Yet prominent Croatian linguists say they are both separate languages (Kapović 2010, Silić 1998). Josep Silić's document is especially important, as he sits on a committee that is in charge of the standardization of the Croatian language (Silić 1998).

Recently a Croatian linguist forwarded a proposal to formally recognize Čakavian as a separate language, but the famous Croatian Slavist Radoslav Katičić argued with him about this and rejected the proposal on political, not linguistic grounds. This debate occurred only in Croatian linguistic circles, and the public knows nothing about it (M. Jembrih 2014).

Despite claims to the contrary, Čakavian is not intelligible with Serbo-Croatian (M. Jembrih 2014).

In fact, Slovene has better intelligibility of Čakavian than Serbo-Croatian does. A number of Croatian linguists including Mitjel Yoshamya have proposed a strong connection between Čakavian and Slovene. Yoshamya based his theory on the study of Archaic Islander Čakavian speakers on the Rab and Krk Islands. This is considered heretical in Croatia for political reasons. Yoshamya was killed during Communist rule in Yugoslavia. There are claims that he was murdered for pursuing his unorthodox linguistic studies.

In the far north of the Čakavian area in Istria where Čakavian and Slovene communities meet, Čakavian and Slovene share some similarities, for instance, final $-m > -n$. These similarities are probably due a lot more to so-called Adriaticisms or an Adriatic Sprachbund than to common genetic heritage. The shared Sprachbund features have the result of making languages not very far from each other in the first place seem even closer than they already are.

Indeed, a study lasting 30 years suggested that Russians and Slovaks understand Čakavian better than Serbo-Croatians do.

The reason for the closeness of Čakavian with Russian is because Čakavian speakers in the islands continued to use OCS in their masses all the way up until the 1970's. This was actually more of a mixture of OCS and Čakavian which was often called "Glagolitic." Due

to modernizing trends by the Vatican, it was suspended it 40 years ago and replaced with Standard Čakavian and Serbo-Croatian.

In recent years, masses have been restarted in its old OCS-Čakavian form on some weekends and church holidays in the northern Islands, in some Kvarner Gulf towns, and also in some Franciscan churches in the interior of Croatia, including three churches in Zagreb. Hence the impact of OCS on Čakavian was considerable. In contrast, OCS had almost no effect on Kaikavian or Štokavian, as these Catholics used the Latin Mass instead.

The following Čakavian word list indicates the similarity of Čakavian with Russian. None of the Čakavian vocabulary below is present in Serbo-Croatian.

In this case, the Russian used is the *Primorskje dialect* of Russian spoken on the White Sea: *bližik* "kinsman", *dosadit* "bore", *dymnik* "chimney", *hlam* "hilltop", *kostje* "skeleton", *krosna* "loom", *krotit* "to drill", *pahat* "swing", *pestat* "destroy", *skocit* "jump", *soha* "fork-rod", *šuica* "left arm", *zaspāt* "to sleep", *babuljki/babulji* "shore-shingle", *kotiga/kotriga* "a dress", *v'vek/vavek* "always", *pavna/plavna* "mud", *revet'/revat* "to roar", *zastruga/struga* "river bed", *zavopet'/zavapit* "exclaim".

The main Čakavian language consists of at least five major dialects.

Ekavian Čakavian is spoken in the east of the Istrian Peninsula, inland around Žminj and Pazin, in the coastal region down to Bakar, on Cres Island, and in the northern part of Lošinj Island (Kalsbeek 1998).

Ekavian Čakavian is mainly distinguished by the presence of the neocircumflex or otherwise unexpected long falling vowel in certain restricted circumstances. All other branches of Čakavian lack the neocircumflex (Kalsbeek 1998).

The Buzet branch has ɛ for the yat (Kalsbeek 1998).

For the the yat vowel, North Čakavian has mostly e. On Krk Island and in two areas in the Croatian Primorje - from Hrejlin south to Novi Vinodolski and Ledenice; and around Grobnik just north of Rijeka - it has either *i* or *e* according to Jakubinskij's rule (Kalsbeek 1998, Jakubinskij 1925).

Ekavian Čakavian traditionally has two branches, Buzet and Northern Čakavian.

Buzet, Buzetski, Buzetska, Bužeština, Buzetska-Gornjomiranski, Gornjomiranski, Upper Mirna Valley Slovene Čakavian is originally an Ikavian-Kaikavian dialect, but it is so heavily Slovenized that it is [fully intelligible](#) with Slovenian, and it now lacks full intelligibility with the main Kaikavian region in Northern Croatia.

It is spoken in the remains of the old former Kaikavian-Čakavian transition zone in Istria in an area around the city of Buzet between Roč and Blatna Vas in the southeast, Draguč in the south, Brkač, Livade, Oprtalj, and Pregara in the west, and the Croatian-Slovenian border in the north and northeast. To the northeast, the line separating Buzet from Slovene is not clear at the moment. The best analysis is that the dialects with stress retractions are Slovene and those without are Buzet. So Slum, Brest and Klenovšac are Slovene and Praproče, Račja Vas, Podgaće, and Lanišće are Buzet (Kalsbeek 1998).

This is best seen as a Čakavian dialect transitional to both Slovene. As mentioned above, Buzet has ɛ for the yat (Kalsbeek 1998).

Speakers refer to themselves and their language as Kaikavian, but more recently, some have started to say their language is actually Slovene.

Buzet has a lot of Northern Čakavian elements in it along with some Slovene (M. Jembrih 2015).

But there are clear Kaikavian traits. This was originally a Kaikavian dialect which got isolated in Istria with the nearest Kaikavian dialects dozens of miles away further east and subsequently came under heavy Čakavian and Slovene influence. This means that at one time, Kaikavian extended all the way to Istria, otherwise there is no explanation for an Old Kaikavian dialect in a region totally isolated away from the rest of Kaikavian (M. Jembrih 2015).

Researchers have classified Buzet as Kaikavian (Rac and Lovrić 2006), connecting it with other Ikavian Kaikavian dialects deep in the Croatian interior such as the Lower Sutla dialect in Zagorje (M. Jembrih 2015). It also has connections with the most ancient forms of Old Croatian such as Gorski Kotar Kaikavian, Bednjanski Kaikavian, and Archaic Islander Čakavian languages such as Gan-Veyã.

Indeed, older Buzet speakers in the Buzet Valley in Upper Raša River region in the villages Mlun, Nugla, etc. still [retain an ancient form of](#)

[the dual](#) which has gone out of almost the entire rest of Macro-Croatian. There is some controversy regarding this Buzet dual. Some think it is occurring due to recent heavy influence of Slovene, which retains a dual. Others say that this is the remains of the Old Croatian dual. The fact that it is only spoken by elders in this isolated valley and that this is such an ancient Croatian dialect implies that the dual is original.

This dialect has a lot of Italian and Istroromanian borrowings. Buzet has very good but not full intelligibility of the North Kaikavian spoken in Northcentral and Northeast Croatia. Buzet intelligibility of North Kaikavian is hampered by the fact that North Kaikavian has many more Germanisms than Buzet does. Buzet may have 75% intelligibility with Northern Kaikavian. MI between Buzet and Slovene is full, with the Littoral dialects being the most similar.

Buzet is similar to the remains of the Ikavian-Ekavian Kaikavian dialect spoken in the former Kaikavian-Čakavian dialect continuum in Central Croatia and is apparently the western edge of this dialect (M. Jembrih 2015). This Ikavian-Ekavian Kaikavian dialect is discussed under the Kaikavian section.

New analyses say that this is a Kaikavian dialect which became Čakavian and recently came under heavy influence of Slovene. Old census publications refer to this population as Slovene. There is still some uncertainty over whether this is a Slovene dialect or a Čakavian dialect. MI is full with Slovene, in particular the Littoral dialects. MI between Buzet and Čakavian is not known. Structurally, the dialect seems more Čakavian than Slovene, but the MI information is still uncertain, so its classification is still somewhat up in the air.

Northern Čakavian or Sjevernočakavski is spoken in Northern and Eastern Istria from the Slovene border to a line that runs from Northern Lošinj to Novi Vinodolski, in Central Istria in Pazin; across the border into Slovenia and across onto the mainland around Rijeka; down the coast to the Novi Vinodolski region; and on Krk and Cres Islands (Kalsbeek 1998).

In Northern Istria it is spoken in Blatna Vas, Brezovica, Kraj Gorenji, Brešca, Grabova, Lovran, Marčelji, Matulji, Mošćenice, Rukavac, Skalnica, Sv. Matej, Volosko, Zvoneće Brkač, Črnica, Donat, Gornja, Donja Nugla, Veprinac, Zatkan, Gradinje, Ivan, Jurišići, Korelići, Krbavšići, Krkuš, Lanišće, Oprtalje, Livade, Mlum Mali, Zagorje,

Marčeniĝla, Drašćići, Oslići, Podgaće, Praproće, Pregara, Račice, Račja Vas, Ročko Polje, Strana, Strped, Mlum Veli, Sv. Jelena, Veliki Mlum, Zamašk, Zrenj, Žonti, Rupa, Lipa, Škalnica, Breza, and Studena. Buje is the far western end of Northern Čakavian (Vermeer 1982).

In Central Istria around Labin it is spoken in Diminići, Rabac Gorenji, Kranjsko Selo, Letajac, Sv. Nedelja Labinska, Polje, Prklog, Rabac Luka, Ripenda-Kosi, Skitača Sv. Lucija na Skitači, Vičani, Ripenda-Kras, Debeljuhi, Kršanci, Orbanici, Vlašići, Žminj, Pazinština Beram, Bortol, Floričići, Gračišće, Heki, Katun Lindarski, Povile, Zagore, Plomin, Staro-Pazinski, Lindar, Novaki Motovunski, Pazin, Pićan, Pilati, Pulići, Tinjan, Trviž, Velanov Breg, Andrevići, Boljunsko Polje, Dolenja Vas, Lupoglav, Mandići, Pas, Semić, and Vranja (Vermeer 1982).

It is spoken across into the mainland in the north in the Hrvatsko Primorje in the old town districts of Rijeka such as Trsat and in Hrejin, Meja-Gaj, Kostrena, Kukuljanovo, Plosna, Podbežice, Ponikve Krasičke, Praputnik, Škrljevo, Urinj, Baštijani, Drastin, Dražice, Ilovik, Donje, Kukuljani, Orehovica, Podhum, Soboli, Trnovica, Kaštav, Marčelji, Praputnjak, Krasica, Zastenice, Grobnik, Jelenje, Čavle, and Cernik Primorski, and Bakar (Vermeer 1982).

To the south in the Novi Vinodolski area, it is spoken in Kraljevica, Crikvenica, Bakarac, Hreljin, Zlobin, Drivenik, Grižane, Jadranovo, Ledenice, Mali Dol, and Selce (Vermeer 1982).

There are a few Middle Čakavian speakers among the the SW Istrian Čakavian speakers in Western Istria around Poreč and in Southwestern Istria around Pula and among the Gorski Kotar Kaikavian speakers in the Delnice area (Vermeer 1982).

All of these dialects except Meja-Gaj and Crikvenica have an *i/e* for the yat instead of the usual *e* (Vermeer 1982).

On the islands it is spoken on Krk Island in Čižići, Dobrinj, Gabonjin, Gostinjac, Hlapa, Sv. Ivan Dobrinjski, Klanice, Kivna, Kras, Polje, Rosopasno, Rudine, Soline, Sužan, Šilo, Tribulje, Sv. Vid Dobrinjski, Županje, Omišalj, Vrbnik, Gornja, Donja, Kozarin, Risika, Paprata, and Punat; on Cres Island in Belej, Beli, Cres, Dragozetići, Filozići, Grmov, Ivanje, Jakov, Lubenice, Martinščica, Miholaščica, Nerezine, Orlec, Osor, Pernat, Predošćica, Punta Križa, Stivan, Ustrine, Valun, Vodice, and Vrana; and in Nerezine and Sv. Jakov on the far northern part of Lošinj Island (Vermeer 1982).

On Krk Island, Omišalj, Vrbnik, Punat, Stara Baška, Baška, Draga Bašćanska, Čižići, Dobrinj, Gabonjin, Gostinjac, Hlapa, Sv. Ivan Dobrinjski, Klanice, Kivna, Kras, Polje, Rosopasno, Rudine, Soline, Sužan, Šilo, Tribulje, Sv. Vid Dobrinjski, and Županje all have *i/e* for the *yat* instead of the expected *e* (Vermeer 1982).

In the town of Opatija, it is still widely spoken by people of all ages. Northern Čakavian is closer to [Kaikavian and Slovene](#) than it is to Štokavian. Northern Čakavian has many Adriatic Venetian loans, as this Venetian language was once widely spoken in this part of Istria and is still spoken in a few places (Kalsbeek 1998).

It is said to be part of Archaic Islander Čakavian due to its nonpalatalization, but as it is fully intelligible with the rest of Čakavian Proper, it is best placed in Čakavian Proper than in Archaic Islander Čakavian.

Rijeka Northern Čakavian has 100,000 speakers in the city of Rijeka, representing half the population. This is the largest concentration of Čakavian speakers in Croatia. 25% of residents are Kaikavian immigrants from inland, and the rest are Štokavian speakers. There are only a few speakers of the Romance language Fiume Venetian, a type of Venetian close to Triestine, remaining. Most remaining speakers are in Italy, where it is also dying out.

[Punta Križa Cres Island Northern Čakavian](#) is spoken in the far south of Cres Island. It resembles the Middle Čakavian dialects of Lopar spoken on Rab Island and Draga Bašćanska spoken on Krk Island. All dialects on Cres Island are Northern Kaikavian.

Nerezine Lošinj Island Northern Čakavian is a Northern Čakavian dialect spoken on the far northern end of Lošinj Island. This is the far southwestern end of the Northern Čakavian zone.

It is dying out because the full-time residents have been leaving the island to look for work, typically to the US, for decades. There is not much of a way to make a living. As of 1989, there were only four full-time families left in Čunski who had been living there for at least two generations. Almost all of the other houses are vacation second homes that are only occupied part of the year. Most of the younger people have a decent understanding of the local dialect, but they cannot be considered fluent. There are a handful of fluent speakers left, all in their 80's.

A few dialects are hard to characterize, as they seem to be transitional between Northern and Middle Čakavian.

Sveti Jakov Lošinj Island Čakavian is spoken in Sveti Jakov, just south of Nerezine, a Northern Čakavian dialect, and just north of Čunski, a Middle Čakavian dialect. Sveti Jakov is hard to characterize. The best description is that it is transitional between Northern Čakavian and Middle Čakavian.

North Čakavian intelligibility of other Slavic lects:

North Čakavian has 95% of Southern Čakavian.

Ekavian-Ikavian Čakavian is represented by a single dialect, Middle Čakavian.

Middle Čakavian or Srednjočakavski, the most widely spoken Čakavian dialect, is Ikavian-Ekavian transitional. It is characterized by an absence of the neocircumflex and a reflex of yat that is either *i* or *e* following Jakubinskij's Rule (Kalsbeek 1998, Jakubinskij 1925). However, in these Ikavian-Ekavian dialects, there are always some exceptions to the *i/e* rule (Houtzagers 2008b).

It is spoken in Central Istria and in Northeastern Istria north to Obrova in Slovenia.

In Central Istria, it is spoken around Sv. Petar u Šumi near Žminj, connected by a small strip including the towns of Milotski Breg and Andretici near Pican east to north of Nedeščina/Sv. Nedelj in the Labinstina area of East Istria north to the area of northeast of Pazin, including the towns of Gradinje, Gologorica, Cervovlje, Novaki Pazinski, and Zarečje. It includes such towns as Cere southeast of Žminj and Polje Čepić north of Nedeščina, and Mrkoč in the center of the area (Kalsbeek 1998).

In Northeastern Istria, it is spoken in the far northeast in Šapjane, Veli Brgud, and Mali Brgud around Mune to a couple of miles across the Slovenian border (Kalsbeek 1998).

The corridor is not contiguous, as it is broken by a Northern Čakavian section between Sv. Petar u Šumi and Polje Čepić. The Istrian speakers and those of the Romanian village Zejane migrated from a homeland in the eastern part of the Lika Highlands (Kalsbeek 1998).

It is also spoken in the islands from the north on Western Krk Island in Njivice and Dubašnica south on the Rab, Lošinj, Pag, Ugljan, Long,

and Dugi Otok Islands all the way to Kornati Island; along the coast south from Donja Draga a mile outside of Rijeka south to Senj; a bit inland at Otočac, Brinje, and Vrbovsko in the Lika Highlands (Vermeer 1982); in Bosiljevo near Severin na Kupi; around Bregana near Samobor; and in Austria, Hungary, Romania, and Slovakia in the northern half of the Burgenland Croatian area (Kalsbeek 1998). Senj no longer speaks Čakavian.

Brinje and Otočac are both inland of Senj, Brinje 12 miles to the west and Otočac 18 miles to the southwest. Vrbovsko is halfway between Rijeka and Karlovac seven miles south of the Slovene border near the town of Skrad. Bosiljevo is 27 miles west of Zagreb on the Slovene border. Bregana is 15 miles west of Zagreb on the Slovene border.

There is a tiny island of ethnic Slovene Middle Čakavian speakers in Slovenia near the Croatian border 2-5 miles west, northeast, and southeast of Podgrad near the city of Ilirska Bistrica in the towns of Golac, Dobro Polje, and Starod. All told, there are only 195 people in these villages in this economically depressed area.

Some villages in Croatia along the Bosnian-Croatian border in Northern Bosnia and Southeastern Croatia beside the Sava River along its northern bank in Croatia at the Vrbas River Estuary - Davor, Orubica, Siće, and Magićmala - speak a Middle Čakavian dialect almost identical to and containing nearly the same lexicon, phonology, and stress as Middle Čakavian spoken in the Kvarner Archipelago (M. Jembrih 2014).

This is far to the east of the present day Čakavian-speaking area. These outlying speakers are possibly the far eastern remnant of the Čakavian area. Alternatively, Čakavian speakers may have migrated east to this area from the Kvarner Archipelago during the Turkish invasions. These speakers have been speaking this archaic Čakavian dialect for 500-700 years.

Ćićarija Čakavian or Srednje-Čakavski Middle Čakavian is a Middle Čakavian dialect spoken on the northern part of the Istrian Peninsula in Slovenia near the Croatian border. It is usually classed as a Slovene dialect transitional to Čakavian., but a closer look shows that it is more Čakavian than anything else. Furthermore, Slovene has only partial intelligibility of Ćićarija, so this can't be a Slovene dialect. It also has some Old Kaikavian elements. The speakers say they speak Kaikavian.

The Slovene elements are obviously areal, but the Kaikavian elements had to have been there from the start (M. Jembrih 2015). This was originally a Kaikavian dialect that has come under heavy Čakavian and Slovene influence, more Čakavian than Slovene. It has only a small number of speakers.

Ćićarija is close to Buzet, although Ćićarija speaks differently from Buzet. Nevertheless, Buzet and Ćićarija are similar structurally. Both are Old Kaikavian dialects spoken in Northern Istria that have come under heavy Čakavian influence and less influence from Slovene. Both groups of speakers call their dialects Kaikavian. However, Buzet has full MI with Slovene, while Slovene only has partial understanding of Ćićarija. Intelligibility with the rest of Čakavian is unknown as in the case of Buzet.

[Draga Bašćanska Krk Middle Čakavian](#) is spoken in Southeastern Krk Island near the famous Gañ-Veyãñ Archaic Islander Čakavian. It is close to the Lopar Middle Čakavian dialect on Rab Island the Punta Križa Northern Čakavian dialect spoken on Cres Island.

Čunski Lošinj Island Middle Čakavian is spoken in the center of Lošinj Island, south of Sveti Jakov. This is at the far northern end of the Middle Čakavian zone.

[Rab Middle Čakavian](#) is spoken throughout the island in small villages such as Barbat, Kampur, and Supetarska Draga. This is very similar to the dialect spoken on Pag Island. It has also retained the Old Croatian perfective imperfect discussed in the Pag section below (Kalsbeek 2008).

[Lopar Rab Middle Čakavian](#) spoken in the far north of Rab Island is different from Rab in the rest of the island. It is more similar to a Middle Čakavian dialect spoken on Krk and a Northern Čakavian dialect spoken on Cres.

[Pag Middle Čakavian](#) is spoken in various places on the island of Pag. It is spoken in the northwest from Luna to Novalja and in neighboring villages. A weaker form of this dialect, the Dinjiška-Povljana dialect, is spoken in the south of the island. Both have retained the odd Old Croatian perfect imperfect as iterative past (Kalsbeek 2008). There are reports that Pag Island is now only half Čakavian-speaking.

Ugljan Island Middle Čakavian is a Middle Čakavian dialect that is nevertheless close to Zadar Southern Čakavian since both have

undergone heavy Štokavization. There were many Italians in Zadar and Ugljan Island, but they left after World War 2. They were then replaced by Štokavian speakers from the hinterlands of Croatia, which began the Štokavization of these Čakavian dialects. Croatian intelligibility of this dialect is quite high due to its extensive Štokavization. Half of Ugljan Island no longer speaks Čakavian.

Southern Čakavian intelligibility of other Slavic lects:

Southern Čakavian has 95% understanding of Northern Čakavian and Kvarner Archipelago Middle Čakavian.

Ikavian Čakavian is spoken in the south and west and a bit in the northwest and northeast of the Istrian Peninsula, on the islands of Brač, Hvar, Vis, Korčula, and Šolta, on the Pelješac Peninsula, on the Dalmatian coast at Zadar, in the outskirts of Split, and inland at the Gacka Valley. There are few speakers left in the city of Split itself, and those that remain are old people living in the Varoš District in the center of the city. About 10% of the population of Split speaks a half-Čakavian mixed dialect.

Ikavian Čakavian has two branches, Southwestern Istrian and Southern Čakavian. Both branches have *i* for the yat vowel (Kalsbeek 1998).

Southwestern Istrian Čakavian, Čakavian-Štokavian or Ikavian Štakavski (SW Istrian Čakavian) is spoken in Southern and Western around Vojnan and Pula, a bit in Northwest Istria, and a small pocket in Northeast Istria around Vodice. It is characterized by the absence of the neocircumflex and an *i* reflex for the yat (Kalsbeek 1998).

In the northeast, it is spoken in the villages of Vodice, Jelovice, Dane, Trstenik, and Raspor. It is also spoken in the four small villages of Brdo, Godjaki, Zagrad, and Črnjehi. This group is divided into two types: a Trstenik type spoken in Trstenik and Raspor and a Dane type spoken in the rest of the villages (Kalsbeek 1998).

In Northwest Istria, it is spoken to the north of and just to the south of the Mirna River east to a bit west of Livade and Motovun and north to the Slovenian border (Kalsbeek 1998).

In the south, it is spoken south of a line extending from Tar to Gordoselo and west of a line extending from west of Oprtalje and Livade south to Motovun, then east to Kašćerga, Kršikla and Gordoselo, then west to Karobja, then south along the Draga River to

Kanfanar, then southeast to Barban and the mouth of the Rasa River on the east coast of Istria. Everything south and west of that line is SW Istrian Čakavian. Really everything south of the Mirna River is SW Istrian Čakavian (Kalsbeek 1998).

However, in Pula, most people now speak a Štokavian dialect called Pula Štokavian. This is the only part of Istria that speaks Štokavian. The rest of the island all speaks Čakavian. This is old Ikavian Čakavian-Štokavian with heavy influence from Neo-Štokavian that retains many Čakavian words and Venetian borrowings. See more under the Štokavian section.

For a long time starting in the 1920's, SW Istrian Čakavian was thought to be a Štokavian or Ikavian Štokavian dialect. A later suggestion was that it was a Štokavian dialect transitional to Čakavian. Even later in the 1960's, a new theory said this was a Čakavian dialect transitional to Štokavian because Čakavian elements predominate.

SW Istrian Čakavian has been described down through time beginning from the earliest classification to the latest as: mixed Čakavian-Štokavian, Štokavian with Čakavian elements, transitional Štokavian-Čakavian, Štokavian with the northernmost area Štokavian-Čakavian, Čakavian-Štokavian with mostly Čakavian characteristics, and Čakavian (Kalsbeek 1998).

Exemplifying this view, in 1964, Mate Hraste said, "These (SW Istrian Čakavian) dialects are not Štokavian nor Štokavian-Čakavian as Ribarić said, but Čakavian-Štokavian because even today Čakavian elements prevail and not Štokavian ones. The only Štokavian dialects in this group are the Premantura dialects in Premantura, Banjole, Vintijan, Vinkuran, and Valdebek, although even in those places there is Čakavian adstrata which was introduced from the hinterland for centuries from the time they arrived until today. The language of all these places even today is mostly Čakavian" (Pliško 2009).

Around Pula are a number of dialects called the Premantura dialects of which the classification is uncertain. They are Lobarika, Muntic, Valtura, Jadreski, Sikici, Skatari, Lijnjan, Medulin, Pomer, Premantura, Banjole, Vinkuran, Vintijan and Valdebek. They are so Štokavianized that controversy about whether they are Štokavian or Čakavian has lingered even after most of SW Istrian Čakavian was finally determined to be a Čakavian dialect, albeit with heavy Štokavian

influence.

Above, Hraste is trying to have it both ways and is dodging the affiliation of the Premantura dialects. He says they are Štokavian, but then he says they are mostly Čakavian. Which is it then?

Since then, [newer research](#) has determined that the Premantura dialects are not Štokavian. Instead they represent a subgroup of SW Istrian Čakavian that has even more Štokavian elements than already heavily-Štokavized SW Istrian Čakavian itself (Mandić 2012).

The Premantura dialects [probably represent](#) the last remnants of the Čakavian-Ikavian Štokavian dialects that the Dalmatian migrants brought with them when they settled the region in the 1500's and 1600's. The rest of SW Istrian Čakavian became much more Čakavized in the centuries that they were present in Istria. This occurred in Pula also, but Pula became mostly Štokavian recently due to Neo-Štokavian influence from the standard language (Mandić 2013).

SW Istrian Čakavian is the largest Čakavian dialect on the southern coast of Istria. It was founded by a group of Čakavian-Ikavian Štokavian speakers who moved to the north to form refugee camps when they were fleeing the Turkish invasions.

There is another theory that these were Dalmatians invited north by Venetian authorities (the area was under Venetian rule) after it had become depopulated due to wars and epidemics. They came from the 1400's-1600's, with most coming in the 1500's. They mostly left from Zadar and Šibenik, which points to a homeland around Central and Northern Dalmatia, a restricted region which explains the homogeneity of SW Istrian Čakavian (Kalsbeek 1998).

This language already had a large Čakavian substrate in Dalmatia, but it subsequently came under further Čakavian influence in Istria. This is an ancient Dalmatian dialect that has come under such heavy Čakavian influence that it is now a Čakavian dialect. It is best seen as a Čakavian dialect transitional to Dalmatian.

The Čakavian elements here are completely different from the native Istrian ones in Northern Čakavian. The theory is that Southwestern Istrian represents a language lost in time, the Čakavian-Štokavian dialect spoken in Dalmatia from 1475-1650. We know that SW Istrian Čakavian has [difficult intelligibility](#) with Middle Čakavian because

Banjole SW Istrian Čakavian cannot understand Sv. Petar u Šumi Middle Čakavian. These two dialects are completely different, and SW Istrian Čakavian is frozen back in time and has a huge layer of Ikavian Štokavian.

However, SW Istrian Čakavian [intelligibility of Southern Čakavian](#) is excellent, so splitting SW Istrian Čakavian off from Čakavian Proper is not justified.

Molise Slavic, an old Dalmatian language with heavy Čakavian substratum, has the same origin as this language. It originated from the same wave of migrants fleeing Turkish attacks in Dalmatia. Most went to Southern and Western Istria, but a few went across the Adriatic to Italy.

Southern Čakavian or Južnočakavski is spoken in Northwest Istria around Poreč, Umag, and Vodice; in Klana, a village near Rijeka; down along the coast in the Čakavian-speaking areas from Senj south to areas near Split, Omiš, Trogir, Zadar, and Biograd and the western half of the Pelješac Peninsula; and on the Čakavian-speaking islands south of Dugi Otok such as the Brač, Hvar, Vis, Korčula, Kornat, Pašman, and Šolta Islands (Kalsbeek 1998). The dialect is dying out in Umag and is no longer spoken in Biograd or Omiš. Šolta, Pašman, and Korčula are now only half-Čakavian.

It is also spoken in the far southern part of the Burgenland Croatia area in Hungary, Slovakia, and Austria (as Burgenland Croatian), and in three villages in the Molise area in Italy (as Molise Slavic) (Kalsbeek 1998).

At the far southern end it merges with Dalmatian around Dubrovnik. It is thought that Southern Čakavian used to extend further inland but was pushed back by expanding Štokavian speakers. The Štokavian influence is so strong in Southern Čakavian that this dialect is considered to be Čakavian-Štokavian transitional.

Both Southern and Southwestern Čakavian are significantly mixed with Štokavian, Southwestern Čakavian in particular.

Southern Čakavian is characterized by an absence of the neocircumflex and an *i* reflex of the yat (Kalsbeek 1998).

Ijekavian Čakavian is the third major branch of Čakavian based on its having the *ie* reflex of the yat vowel (Kalsbeek 1998).

Southeastern Čakavian, Ijekavian Čakavian or Lastovian is spoken at the far southern end of the Čakavian language area on Lastovo Island, in Janjina on the Pelješac Peninsula. There is a small enclave of Southeastern Čakavian speakers far to the south of the main speaker area at Bigova in Kotor Bay on the coast of Montenegro. There are also some speakers in the United States in Los Angeles and New Orleans and in Canada in Vancouver.

Southeastern Čakavian is only spoken in a tiny area and has few speakers. This dialect used to be much more widespread, but it has lost a lot of its territory and has been pushed into a few remaining refugia. It has similarities to Southern Čakavian.

Čakavian intelligibility of other Slavic lects:

Čakavian intelligibility calculation was good, with a number of good Čakavian informants, in particular a few excellent personal informants, including linguists.

Čakavian has mostly good to excellent intelligibility within itself. Southern Čakavian has 95% intelligibility of Northern Čakavian and Kvarner Archipelago Middle Čakavian (Kolanović 2015), and speakers of *Labin Northern Čakavian* in the far north say they can understand the speech of the southern islands very well (M. Jembrih 2014).

Southwestern Istrian Čakavian spoken along the southern coast of Istria cannot understand the Middle Čakavian spoken in the interior.

Čakavian may have good intelligibility of Kaikavian at 80-85% (M. Jembrih 2014). Reports of [full intelligibility](#) between the two languages are not correct.

Čakavian intelligibility of Slovene is 40%. Čakavian is said to be fully intelligible with Slovene. This cannot be the case, but intelligibility may be fairly good. Čakavian and Slovene both share a lot of archaic Slavic vocabulary. Nevertheless, Čakavian's intelligibility of Slovene is worse than that of Kaikavian or Serbo-Croatian.

Čakavian intelligibility of Bednjanski Kaikavian is 40%.

Čakavian has 40% intelligibility of Slovak, a remarkably high number.

Čakavian Proper has 32% intelligibility of Archaic Islander Čakavian, which proves that these are absolutely separate languages.

Čakavian has 29% intelligibility of the highly archaic **Rhodope Mountains Pomak**. It is interesting that three Serbo-Croatian

informants could not understand one word of Rhodope Mountains Pomak, and yet two types of Čakavian scored much better.

This is very interesting because one would not expect Čakavian and **Pomak** to understand each other well. But Čakavian is one of the most archaic Slavic languages, as archaic as Russian and Bulgarian. Rhodope Mountains Pomak is also highly archaic. Perhaps Pomak and Čakavian are the remains of a highly archaic form of Balkan Slavic that was once much more widely spoken in the Balkan region long ago, possibly over 1,100 years ago. We could perhaps call this Proto-South Slavic or Proto-Mediterranean Slavic. Perhaps this was the remains of a seafaring koine used by fisherman and others around the Adriatic and Black Seas.

Čakavian intelligibility of Bulgarian is poor at 20%.

Čakavian has drastically poor intelligibility of Molise Croatian at 5%. This is interesting because this a Čakavian dialect that moved to Italy and came under heavy Italian influence for 500 years. It's remarkable that the intelligibility deteriorated so far, but perhaps the Italian contact was extreme.

Archaic Islander Čakavian

Archaic Islander Čakavian, Adriatic Tsakavian, or Boduli is the oldest form of Čakavian in the Čakavian macrolanguage. It is no doubt not even a single language – instead it is a macrolanguage containing at least five languages and possibly more. It is very archaic and has a lot of non-Slavic vocabulary. The phonology and even morphology and syntax is quite different. One dialect has only 4% Serbo-Croatian words, the rest being Romance loans and 25% being from an old substrate, mostly Indo-Iranian.

One of the major differences between this and the rest of Čakavian is that this form is nonpalatalizing. In other words, it has alveolar *c* or *ts* where other forms have the more backed and half-palatal or “palatalizing” *ž, š č* alveopalatals. In most places, these palatals have been replaced with sibilants such as *s, ts,* and *z*. The classic Slavic

palatals are *č, ć, š, ž, dž, dj, lj, and nj*.

Of all of the forms of Čakavian, this is the hardest to understand. Indeed, other Čakavian speakers often can barely understand it. It is dying out everywhere it is spoken.

This language is spoken almost exclusively in Croatia except for a few speakers in North America. In Croatia, it is spoken on the offshore islands including the Kvarner Archipelago, in Istria, and in a few places on the far northwestern coast of Croatia east of the Kvarners.

It is spoken in the towns of Batomál, Šuráje, and Yurandvôr on the southern shore of Krk Island; in the city of Rab on Rab Island; in the coastal village of Novi Vinodolski on the coast opposite Krk Island; in the village of Brseč on the Eastern Istrian coast a couple miles north of Cres Island; in the villages of Davor, Orubica, Siće, and Magićmala in Southeastern Croatia on the border with Northern Bosnia along the Sava River at the Vrbas River Estuary; and in New Orleans, Louisiana, and Vancouver, Canada.

Brač Archaic Islander Čakavian is still spoken to the north and west of the island. There are still entire villages speaking this language in Dol, which is called the Dolici dialect. This dialect cannot be understood at all by Štokavian speakers from as close as Split, only 10 miles away on the coast. Even other Brač islanders can barely make sense of it. MI with the rest to the rest of Archaic Islander Čakavian is not known. Romance borrowings are probably 25-30% of the vocabulary here while Slavic makes up probably 65-70%, but we don't know for sure, as we do not yet have any dictionaries of this speech.

Gan-Veyãň Čakavian is an Archaic Islander Čakavian language that is one of the most archaic of all. Oxytonic stress is present in 50-75% of the lexicon, and the rest of the words have penultimate stress.

It is spoken in the towns of Batomál, Šuráje, and Yurandvôr on the southern shore of Krk Island in Croatia and in New Orleans, Louisiana and Vancouver, Canada. The speaker community is [very small](#): there are only 12 speakers left, all elderly men, either fishermen or shepherds, six on Krk and another six in North America. However, there are apparently still middle-aged speakers of this language, so the situation may not be quite as dire as above.

The Gan-Veyãň spoken in North America has retained much of its purity. Until recently it was spoken by 350 speakers in six villages on

Krk and Rab Islands, but most of them were massacred in foibes after World War 2 by Communist partisans.

From the 500's to the the 1500's, it was [spoken](#) in the Medieval city of Corinthia which was located where Baška is today.

Some reports indicate that Gan-Veyãn is now spoken only in Batomál and Šuráje by a half dozen elderly speakers, but in 2010, there were [still elderly speakers in Yurandvôr](#), so the situation is confused.

It is a very archaic mix of Čakavian, Liburnian, and Early Slavic with Romance loanwords. The [vocabulary](#) is 40% Slavic, 34% Romance, and 23% Indo-Iranian and Mesopotamian archaisms of the Old Orient. The original Croatian language was probably an old Indo-Iranian language from the Orient that was replaced by Slavic immigrants in the 600's. Most accents are terminal.

Like Kyrška Bešeda, Gan-Veyãn is one of the last remaining Croatian languages to retain the dual number. However, this dual like the paucal in other Slavic languages, extends from 2-4.

It is unintelligible to any other Slavic language – Kurdish and Zazaki speakers actually have more intelligibility of this language than Slavs do. Most of the important work done on this language was done by two Croatian anthropologists, the husband and wife team of Mitjel and Zyelimer Yoshamy and Dr. Andria Zeljko Lovric, who continued the work of his father Michael Lovric.

Apparently they all worked together on this dictionary, a 1,222 page monograph on this language that was published by a respectable anthropological research institute in Zagreb. The tome includes 23,000 words from Baška and 5,000 more from Novi Vinodolski and Rab Island.

Nevertheless, a number of Croatians make the odd claim that these researchers made up the story of this language out of whole cloth. This seems odd as the researchers assembled a [monstrosity](#): a 1,121 page book which included a 23,000 word dictionary of the language, a list of ten detailed myths and tales, all sorts of cultural knowledge, 123 unique names of celestial stars that fishers and navigators used to steer by at night, 68 unique words for different types of algae, and the possibility of an ancient matriarchal society.

The roots of this language appear to be in the ancient Liburnian Indo-European language. It seems quite odd that respectable scientists

would make up 23,000 strange words for a dictionary complete with etymologies linking them with other languages, an assortment of myths and tales, abundant ancient cultural knowledge, over 100 unique names of stars, and lists of odd villages speaking this language in Southeastern Croatia, New Orleans and Vancouver out of their heads. They would have to have been the greatest fantasists who ever lived.

In fact other Croatian linguists acknowledge the existence of this language, but they describe as an Old Čakavian dialect with a lot of loans from the extinct Indo-European Illyrian Venetic language and the extinct Macro-Romanian Dalmatian language. Which is not far from what the Yoshamyas came up with, as Venetic was related to Liburnian. Scientific fraud is a very serious charge, and those making it ought to back up their serious accusations with good evidence.

With further investigation, a ready explanation of this mystery is apparent. Yoshamya was killed during the Communist government. The government had ferociously opposed his theories as "fascist propaganda," preferring instead a myth of Yugoslav Pan-Slavism known as Vukovianism. Under the regime, all talk of an ancient Croatian language or heavy Romance influence was attacked as either fascist (Italian imperialism) or Croatian nationalist. Instead Standard Serbo-Croatian, to which many new Turkicisms and Balkanisms were added, was aggressively imposed.

As people supporting this theory were attacked so viciously that at least one of them was actually murdered for his scholarship, it seems clear that a lie stating that Yoshamya made up this whole language and all of the rest of the associated knowledge out of his head was created by the Communist regime in order to disparage him. The lie apparently exists to this day, and there are still Croats who mock this concept and laughingly point out that he made up this whole fake language.

Websites point out that most linguists doubt that it exists, however, a dozen living Croatian linguists say that this language actually exists. One problem is that the Gan-Veyãñ Theory has been taken up by Croatian nationalists, who often are not very pleasant people. In fact it was a favorite of the fascist Ustasha regime, and people promoting the theory today are sometimes White Supremacists. The association of this theory with fascism and racism has no doubt led to its

disparagement and the denial of its truthfulness.

Gan-Veyãñ is dying out and being replaced with Croatian and Standard Čakavian, and the more pure form is spoken mostly by older men. Furthermore, existing Gan-Veyãñ idiolects are being heavily affected by Standard Čakavian and Croatian. It certainly has poor intelligibility with the rest of Čakavian, as Gan-Veyãñ speakers learn Standard Čakavian as a foreign language.

Brayska Besyda Archaic Islander Čakavian is a very archaic little known Archaic Islander Čakavian dialect spoken on the Istrian Peninsula. It is similar to Gen-Veyan. It may be spoken in the tiny village of Brseč on the eastern coast of Istria. Whether this is a separate language or a dialect of Gan-Veyãñ is not known.

Rab Archaic Islander Čakavian is an Archaic Islander Čakavian language spoken by a few speakers in the city of Rab on Rab Island. Rab is said to be the most similar to the Krk Island language. There are still a tiny number of elderly speakers left. 25% of the lexicon is from Ancient Greek and Latin. Despite the similarity with Krk speech, this language still appears to be very different from Gen-Veyan such that MI is surely not full. This language has retained the odd perfective imperfect from Old Croatian to refer to iterative events in the past (Kalsbeek 2008).

Pag Archaic Islander Čakavian is a very well-preserved Archaic Islander Čakavian lect spoken in the center of the city of Pag and the neighboring village of Košljun. It is the subject of a recent extensive dictionary (Kustić 2003). It has retained the same Old Croatian perfective imperfect that the Rab language did (Kalsbeek 2008). MI with the nearby Rab or with the rest of Archaic Islander Čakavian is not known.

Susak Island Archaic Islander Čakavian is so odd that no other Croats can understand it. It is a blend of Old Croatian, French, German, and Italian. Susak was sufficiently isolated from the mainland and even other islands that it developed in isolation for an incredible 1,000 years without linguistic interference from other lects, resulting in this very strange language. It is spoken on the island and by the older generations of people who left the island. The best speech is preserved in people over 60. This language has also retained the same odd Old Croatian perfective imperfect iterative past that Rab and Pag have (Kalsbeek 2008).

Hoboken Susak Island Archaic Islander Čakavian is a form of Susak that is spoken in Hoboken, New Jersey, for three generations. The population of Susaks here is 1,500 (Filipovic 1996).

The group, who refer to themselves as Sušani, first arrived from Susak Island around the turn of the century when most other Croatian immigrants arrived in the US. The first generation never learned to speak English. The women had learned Italian on Susak, and Italians who were recent immigrants and spoke Italian ran many of the local stores, so the women were able to do their shopping at these stores. The men spoke a mixture of Slavic languages which served as a koine for Slavic workers at a local factory where they worked. At home everyone spoke Susak (Filipovic 1996).

Their children, the second generation, learned Susak at home and became full Susak-English bilinguals. As their parents had never learned to speak English, the second generation had a reason to keep speaking Susak because that was the only way they could communicate with their parents (Filipovic 1996).

The children of the second generation, the third generation, as a rule were left with the grandparents during the day while the parents worked. The grandparents do not know English, so they speak Susak to the children. In this way, the third generation learned Susak well too (Filipovic 1996).

The most important factor in maintaining the language has been the family, in particular mothers. There are two other factors. The first is that they all love dancing Croatian folk dances and singing Susak songs. The second is that every Sunday from September-June, the entire Hoboken community has a get-together at a local club with Susak is spoken. The addition of the two cultural factors in addition to the family have helped cement the maintenance of Susak in Hoboken (Filipovic 1996).

There is no data available on MI between Hoboken Susak and Susak back in Croatia.

Kyrška Bešeda Archaic Islander Čakavian is an Archaic Islander Čakavian language spoken in Rupe, Antóvo, and Bašúnje on the southern slopes of Mount Bitoraj around the Croatian coastal Novi Vinodolski in Vinodol Valley, between Fužine and Crikvenica (population 3,500) across from Krk Island. This language and Gan-

Veyǎn are the only surviving Croatian languages that have retained the dual. This dual covers the numbers 2-4 like in the paucal in other Slavic languages. This language appears to be dramatically different from Gan-Veyǎn, and MI between them is probably not full. In fact, this language is extremely diverse within itself and may not even be intelligible within itself.

Komyzjonski Archaic Islander Čakavian is a very archaic Archaic Islander Čakavian language spoken on the western edge of Vis Island in Croatia near the small coastal village of Komiža, population 1,500. It is nearly but not quite Slavic, mostly a mix of Čakavian and Ukrainian with 1/3 Romance loans. It is not intelligible to Serbo-Croatian and most other Slavic languages, although Eastern Slavic can understand some of it. It has mostly terminal accents.

Komyzjonski is nearly extinct and is only spoken in a few villages. This language, with its Ukrainian base and Romance loans, cannot possibly be intelligible with any other Archaic Islander speech.

San Pedro Komyzjonski Archaic Islander Čakavian is a dialect of this language spoken by Croatian immigrants in San Pedro, California who work in the fishing industry. Most arrived between 1920-1940. The fishing industry is huge on Vis. At one point, every large European sardine company had a packing plant in Komiža. It is said that there are more Vis Islanders from Komiža in San Pedro than in Komiža itself – there may have been 2-5,000 Komižans in San Pedro as of 1972 (Albin and Alexander 1972).

MI between this dialect and the Komyzjonski Archaic Islander Čakavian spoken back home on Vis Island is not known, but the two lects have been separated for 80-100 years.

Korzulot Archaic Islander Čakavian is an Archaic Islander Čakavian language spoken on the island of Korzula by a few speakers. The language is little known. This seems to be a newer version of these Slavic-Romance lects. It is a mix of Čakavian and Italian and Venetian. Communist partisans massacred many of its speakers in foibes on the islands of Korzula, the island of Brač, and other places. It is now known only from Korzula. This also seems as if it is probably a separate language.

Archaic Islander Čakavian: *Oral intelligibility:* 98% of Old Čakavian†* and Old Church Slavonic†* and 71% of Rhodope Mountains Pomak*.

Archaic Islander Čakavian intelligibility of other Slavic lects:

There was only one good informant for Archaic Islander Čakavian, and much of his work was in Serbo-Croatian, which I don't speak, hence was quite hard for me to read.

Archaic Islander Čakavian has good comprehension of Rhodope Mountains Pomak at 71%*. This is because these are the two most ancient lects in South Slavic, perhaps living fossils of Proto-South Slavic, Early South Slavic, or an ancient mariners' koine used by fishers and sailors on the Adriatic and Black Seas.

Intelligibility testing has shown that Archaic Islander Čakavian has full intelligibility of the extinct Old Čakavian at 98%*. This is because Archaic Islander Čakavian is a language frozen far back in time, and in a sense it is Old Čakavian resurrected and living today amongst us.

Further intelligibility testing has shown Archaic Islander Čakavian has full intelligibility of Old Church Slavonic at 98%*. It is uncertain why this is, but OCS is also an ancient South Slavic language frozen far back in time to the era of Early South Slavic. Further, services were held in OCS in the islands until very recently, so we may be dealing with some bilingual learning here.

Molise Croatian

Molise Croatian, Molise Slavic, Slavomolisano or

Moliškohrvatski is an archaic Dalmatian language spoken in three small villages in Italy, Acquaviva Collecroce and two other villages, Motemiro, and San Felice. A different dialect is spoken in each village. The majority of the speakers are in Acquaviva Collecroce, while Motemiro has fewer speakers who speak a more conservative dialect, and San Felice has only a few older speakers left (Breu 2009).

Despite a lot of commonality between the dialects, the differences between them are significant (Pourheidari 2006). Nevertheless, intelligibility between all three dialects is excellent (Breu 2015).

Yet there are some differences. It is due to these differences that a koine is currently under development (Pourheidari 2006).

These Croatians came from the area of the Cetina and Neretva Rivers in Bosnia Herzegovina.

They arrived in the 1400's (Breu 2009). They were fleeing Turkic invasions and also trying to repopulate Molise which had been devastated and depopulated by earthquakes and epidemics.

Molise Slavic is an Dalmatian language which is in a lot of ways similar to what is now spoken as Southern Čakavian Ikavian on the islands of Croatia. There is also a marked resemblance to Southwestern Istrian.

Nevertheless, Molise Slavic is obviously Štokavian in terms of lexicon, grammar, and phonology, and even Čakavian loans are quite rare in Molise Slavic (Breu 2015).

It has undergone heavy influence from Molise Neapolitan Italian and Standard Italian in the last 150 years (Breu 2009).

Molise Slavic intelligibility of other Slavic lects:

Molise Slavic intelligibility queries benefited from the presence of a linguist expert in the language.

Molise Slavic has 12% intelligibility of Serbo-Croatian.

Serbo-Croatian is almost completely unintelligible to Molise Slavic speakers, and Croatians visiting these villages need to speak Italian if they wish to be understood well (Breu 2015). Breu declined to give an exact figure, but descriptions of this nature typically add up to perhaps 5-20% intelligibility.

Molise Slavic speakers with some training in Serbo-Croatian may understand up to 60% of it, and some who have studied at Zagreb University have full understanding of Serbo-Croatian (Breu 2015).

Molise Slavic has only 5% intelligibility of Čakavian.

Although Molise Slavic speakers have no contact with Čakavian since Croatians visiting the towns use Serbo-Croatian, Molise Slavic's understanding of Čakavian would surely be even worse than their understanding of Štokavian (Breu 2015). Based on this description, an intelligibility of 5% is a reasonable figure.

Burgenland Croatian

Burgenland Croatian is an outlying Croatian language spoken to the

north and east of Croatia. In truth, it is not even a single dialect but instead is a collection of five different dialects that often vary widely and it is not a single language either as it may be more than one language. It is spoken in Austria, Czechia, Slovakia, and Hungary. The standard language is a dialect of Serbo-Croatian, most of the dialects are related to and possibly a part of the Old Čakavian language, but some are Old Kaikavian and others are Old Štokavian. MI figures for the dialects are uncertain.

It is spoken in 65 settlements in the Burgenland region of Austria, 14 settlements in Győr-Moson-Sopron and Vas counties of Hungary, and 5 settlements in the Bratislava district of Slovakia (Paukowits 2014).

Burgenland Croatian is mostly an old Čakavian dialect resulting from a movement of Čakavian speakers fleeing north in the 16th Century to escape the Turkish Wars, part of the same movement that went to Italy and created the Molise Slavic language. The group fled into what was at the time the Austro-Hungarian Empire but is now divided into three separate countries.

The first wave came in 1515 (Pawischitz 2014).

The group included speakers of the Štokavian, Čakavian and Kaikavian languages, but most of the migrants were Čakavian speakers from the Croatian coast (Pawischitz 2014).

The dialect now has many German, Slovak, and Hungarian loans and is highly archaic (Gradišćansko-Hrvatski Centar 2011). It has also come under serious Slovene influence. Structurally, this is an archaic Čakavian Croatian dialect with Slovene case declensions as suffixes. In fact, much of the morphology looks Slovene. The vocabulary is a mix of Slovene and Croatian.

A standardization process was begun 25 years ago. The standard language was based on Čakavian because the Čakavian dialects have the most speakers (Pawischitz 2014).

Burgenland Croatian is now experiencing language death (Pawischitz 2014).

However, the older generation still speaks a relatively full dialect. People over 70 were raised speaking Burgenland Croatian as a mother tongue, were educated in Burgenland Croatian, married other Burgenland Croats, and spoke to their children, families, relations and friends in Burgenland Croatian. They have a very positive attitude

towards the dialect (Pawischitz 2014).

Many of the younger generation, including people about 30-35 years old, were either brought up speaking Burgenland Croatian or they learned it as teenagers and adults. Some are currently speaking to their own children in Burgenland Croatian, so children are still being raised speaking the dialect. The speech of the younger generation shows clear signs of language death, and they can no longer fully express themselves in the dialect. They can only talk about relatively simple things. When they have to talk about something more complex, they resort to paraphrase (Pawischitz 2014).

In the past 30 years, Burgenland Croatian has come to be seen by speakers as an old-fashioned and outdated dialect. A recent change in favor of language revival has only occurred in the past decade (Pawischitz 2014).

Quite a few Burgenland Croats have immigrated to the US, many arriving from 1910-1920.

Most live in places like Pennsylvania, New York and the city of Chicago. A number of them still speak the dialect, and having been isolated from the factors that caused the decline of the dialect in the homeland, still speak a quite pure form of the dialect similar to that spoken in Austria 90-100 years ago (Paukowits 2014).

The number of speakers has been declining from 44,753 in 1920 to 17,730 in 2001 (Pawischitz 2014), a 60% decline over 80 years. The number of speakers has been declining by an average of 7.5% per decade for the last 90 years.

Ethnologue reports that Burgenland Croatian has difficult intelligibility with Standard Croatian. There are other reports that Burgenland Croatian has difficult intelligibility with Standard Croatian (Strazny 2013, p. 105).

A Standard Burgenland Croatian has been created that is used in newscasts and the like. In contrast to reports that Burgenland Croatian has difficult intelligibility with Standard Croatian, all three Croatian informants in my survey who listened to Burgenland Croatian reported very high intelligibility of Standard Burgenland Croatian.

The standard language is a much watered-down version of the authentic local dialects and has been standardized to mirror Standard Croatian, in part to rid it of Germanisms. The purpose was to make

the standard language intelligible to all speakers (Dressler and de Cillia 2006).

The result was the creation of a standard language that almost nobody speaks. The spoken dialects are now very different from the standard language (Pawischitz 2014).

The Štokavization of this dialect could also be due to history. Although Burgenland Croatian started out as an archaic Čakavian dialect, several waves of Štokavization have occurred since the 1500's as new waves of settlers moved into the area from Croatia.

Intelligibility among the various Burgenland Croatian dialects, especially the Kaikavian ones versus the Čakavian ones, is not known.

Different dialects are spoken in different parts of Austria. In the south of Burgenland, Croats speak either an *Old Štokavian* or an *Old Southern Čakavian* dialect; in the middle, they speak an *Old Middle Čakavian* dialect with heavy *Old Kaikavian* influence; and in the far north, they speak an Old Middle Čakavian dialect.

The Kaikavian and Old Štokavian dialects show strong Čakavian influence (Pawischitz 2014).

There are five main types of dialects from north to south: relatively pure Old Middle Čakavian, Old Middle Čakavian with Old Kaikavian influences, Old Kaikavian, Old Štokavian Ikavian Šćakavian, and Old Southern Čakavian.

The Grob, Devínska Nova Ves, Hac and Poljan dialects of the far northern end of Burgenland Croatian are relatively pure Old Middle Čakavian dialects that lack significant Old Kaikavian influences.

Grob Burgenland Croatian of Chorvátsky Grob in Southwestern Slovakia is an unclassified Old Middle Čakavian dialect (Houtzagers 1999). The name Chorvátsky Grob is the Croatian name for the town – the actual Slovak name is different. Chorvátsky Grob means “Croatian graves” or the graves of the Croats. The reference is thought to be to Croatian soldiers, but no one quite knows who these soldiers were. Perhaps it is a reference to the original settlers who were fleeing invasions and massacres of the Turks.

It was originally spoken in between the Štoj, Dolinci, Vlah, and Southern Čakavian speakers, which puts the Štoj speakers along the Una River on the Bosnian-Croatian border around Hrvatska Kostajnica

(Houtzagers 2008a). Grob is definitely going extinct. Old people still speak a pure dialect, but most young people cannot speak the language and do not wish to (Čagalj et al 2014).

There are only 31 speakers of the dialect, less than 1% of the population of the town. However, that figure is a very healthy 38% of the Croatian population of the town.

The dialects in Slovakia are frozen in time and have not been updated to keep up with the modern world. For instance, they have a word for "carriage" but no word for "car." That's how behind the times they are. There are no Croatian-medium schools in Slovakia, but there are four schools that offer classes in Standard Croatian. Speakers are mostly indifferent to the fate of their dialect. Both dialects have been under heavy influence of German, more so than Hungarian. They came under significant Slovak influence only in the 20th Century (Čagalj et al 2014).

Devínska Nova Ves Burgenland Croatian is an unclassified Old Middle Čakavian dialect spoken in Southwestern Slovakia to the west of the Grob dialect (Houtzagers 1999). Devínska Nova Ves came from the same region as the Hac and Poljanci southwest of Karlovac between Tounj and Slunj about nine miles southeast of Zagorje (Houtzagers 2008a). Čakavian is still spoken 11 miles northeast of here.

Devínska Nova Ves still has many active speakers of either the dialect or Standard Croatian, which most Burgenland Croatian speakers also learn. Use of the dialect depends more on membership in churches and cultural clubs than anything else. There are 67 active speakers here, but that is less than .5% of the population. However, it represents 28% of the Slovaks in the town, a more hopeful figure. In the youngest generation, there are only two families that have speakers. People here are very pessimistic about the language and feel it is doomed (Čagalj et al 2014).

Koljnof Burgenland Croatian is an unclassified Old Middle Čakavian dialect spoken in Hungary next to the Hidegség and Fertőhomok Old Kaikavian dialects. It is intermediate between the Poljanci and Dolinci groups. It resembles Poljanci more than Dolinci. Geographically, it is closest to the Dolinci group, but it is very close to the Poljanci group also. It also has some Old Kaikavian influences from Hidegség and Fertőhomok. It is still widely spoken in the town by all generations and in all circumstances and is not going extinct any time soon

(Houtzagers 2008b).

Croatian is taught at the local school and is used in church. Interestingly the Standard Croatian used at church and taught at school is having no effect whatsoever on Koljnof. Speakers understand other varieties of Burgenland Croatian well, probably from exposure, they borrow widely from them to find new words to use to express themselves better. Nevertheless, the borrowings are only used opportunistically on an idiolect level and speakers are well aware of what is a Koljnof word and what is a word from another Burgenland variety (Houtzagers 2008b).

Hac or Hacı Burgenland Croatian near Neusiedl is an Old Middle Čakavian dialect (Houtzagers 1999). The Hacı originally came from the area southwest of Karlovac, an area southeast of Tounj, Ogulin, and Oštarija (Houtzagers 2008a) where a Kaikavian dialect transitional to Čakavian called Ikavian-Ekavian Kaikavian is still spoken. This is the last remains of an old Čakavian-Kaikavian transition zone. Of course this area is no longer Čakavian speaking.

Poljan or Poljanci Burgenland Croatian near Lake Neusiedl is an Old Middle Čakavian dialect (Houtzagers 1999). The Poljanci came from the same region as the Hacı southwest of Karlovac (Houtzagers 2008a).

In the center, the Dolinci and Weingraben dialects are Old Middle Čakavian dialects with significant Old Kaikavian influences.

Dolinci Burgenland Croatian of Unterpullendorf, Frankenau, and Kleinmutschen is an Old Middle Čakavian dialect with Old Kaikavian influences (Houtzagers 1999). The Unterpullendorf dialect cannot be understood at all by Croatian or Kaikavian speakers (M. Jembrih 2015). The Dolinci originally came from the area of Glina and Topusko 20 miles southeast of Sisak and 11 miles north of the Bosnian border (Houtzagers 2008a). The area is no longer Čakavian-speaking.

Weingraben Burgenland Croatian is an unclassified Old Middle Čakavian dialect with Old Kaikavian influences spoken a bit to the west of the Dolinci dialect (Houtzagers 1999). It was probably spoken 14 miles southwest of Sisak along the Kupa River near Slana (Houtzagers 2008a).

In Hungary, the only two Burgenland Croatian Old Kaikavian dialects are spoken.

Hidegség/Fertőhomok Burgenland Croatian is an Old Kaikavian dialect spoken in two towns on far Western Hungary. They originally came from Velika and Meduric in Slavonia at the southeastern end of the Kaikavian zone. The dialect is also mixed with Medjimurje Kaikavian. It's not clear how this mixing occurred. Velika and Meduric are no longer Kaikavian speaking today. This dialect lacks the typical neocircumflex accent present in Kaikavian Proper (Houtzagers 2008a).

It is full of loans, mostly Hungarian. There are also many loans from German and nearby Burgenland Čakavian dialects. It is presently going extinct, as all of the fluent speakers are over 70 years old. The dialect is under extreme pressure from Hungarian and this is what is causing its extinction (Houtzagers 1999).

In between the central dialects and the Old Southern Čakavian dialects of the far south are the Vlah and Štoj dialects, Old Štokavian dialects of the Ikavian Šćakavian type with Slavonian Štokavian features that are transitional to Old Middle Čakavian.

Vlah Burgenland Croatian of Weiden bei Rechnitz, Zuberbach, Althodis, Schandorf, Dürnbach, and Allersdorf is an Old Štokavian Šćakavian Ikavian with Slavonian Štokavian features. It is transitional to Old Middle Čakavian (Houtzagers 1999). Vlah was originally spoken to the south of the Štoj dialect in Bosnia about five miles east of Kosarska Dubica just south of the border near the confluence of the Una and the Sava (Houtzagers 2008a).

Štoj Burgenland Croatian is another Old Štokavian Ikavian Šćakavian transitional to Čakavian similar to the Vlah dialect. It was spoken along the Una River along the Bosnian-Croatian border in an area just northwest of Kozarska Dubica (Houtzagers 2008a).

At the far southern end of the Burgenland Croatian are Hackerburg, Stinatz, and other Old Southern Čakavian dialects.

Southern Čakavian Burgenland Croatian is spoken at the far southern end of the Burgenland Croatian region. It is based on an Old Southern Čakavian dialect (Houtzagers 1999). This Old Southern Čakavian was spoken along the Bosnian-Croatian border from Bihac to Hrvatska Kostanjica along the Una River (Houtzagers 2008a). This area is of course no longer Čakavian-speaking. However, there are still Southern Čakavian speakers 60 miles to the east in the Davor-Srbac region

along the Sava River 15 miles southeast of Nova Gradiska. They have been there for 500-700 years.

Hackerburg Burgenland Croatian is an Old Southern Čakavian dialect (Houtzagers 2014) spoken right next to Stinatz. It is spoken in between Vlah and the main Old Southern Čakavian dialect area.

Stinatz Burgenland Croatian is an Old Southern Čakavian spoken in between Vlah and the principal Old Southern Čakavian dialect area (Houtzagers 2014).

Burgenland Croatian intelligibility of other Slavic lects:

An absolute absence of Burgenland Croatian informants made the question of Burgenland Croatian very difficult.

Burgenland Croatian has full intelligibility of Standard Croatian, as their standard language is based on the Croatian standard.

Burgenland Croatian dialects may have good intelligibility with Čakavian. Reports indicate that Burgenland Croatian speakers can communicate well with Dalmatian Čakavian speakers, whereas Standard Croatian speakers cannot communicate well with them at all, and sometimes resort to German in order to speak to them (M. Jembrih 2015). While Standard Burgenland Croatian is a dialect of Croatian, the actual spoken vernaculars are mostly dialects of the Čakavian language. However, not all Burgenland Croatian speakers speak Čakavian. Some speak Kaikavian and Old Štokavian Šćakavian Ikavian.

Serbian

Serbian is a macrolanguage made up to two languages: *Serbian* and *Torlakian*.

Serbian is simply the same Serbo-Croatian language that is also spoken in Croatia, Montenegro and Bosnia/Herzegovina. It forms a single tongue and not separate languages as many insist. The claim for separate languages is based more on politics than on linguistic science.

Standard Serbo-Croatian was standardized based on the Ijekavian Štokavian dialects of Eastern Herzegovina. Yet the new Standard Serbian is based on the Ekavian speech that is more common in

Serbia. Nevertheless, the new Standard Serbian is largely ignored in schools and other institutions in favor of Standard Serbo-Croatian. However, others say that forcing Ijekavian Štokavian on Ekavian Štokavian speakers has not worked well and contributes to a lot of what critics say is functional illiteracy in Serbia.

It is said that there is not much difference in Serbian dialects other than the yat, but this is not the case. The Ijekavian Štokavian spoken near the Croatian border and the Ekavian Štokavian spoken in Western Serbia are very close to the standard language, but other dialects spoken in Serbia are not.

The earliest Serbian writing is in Slavic Church Language and dates from the 1200's. Most of these have an ecclesiastical/panegyric nature. The Turkish conquest of Serbia which began in the Battle of Kosovo in 1389 and ended with the final conquest in 1459 put an end of intellectual and literary life in Serbia (Author Collective 1885-1892).

The center of Serbo-Croatian scholarship moved to the state of Ragusa in Dalmatia, then under the control of Italians. This period of Dalmatian literature from 1500-1700 was originally in Croatian, but after a while, the center of gravity shifted to Serbian. On April 7, 1667, an earthquake destroyed Ragusa. After that it was no longer a prosperous town. With that decline, Ragusan literature also deteriorated, and in the 1700's, it was only known in an impoverished form (Author Collective 1885-1892).

Most Serbian writing pre-1850 which was written in a combination of Old Church Slavonic and local Serbian dialects (Author Collective 1885-1892). Even now, Serbian has many Old Church Slavonic words that are either archaic or lack cognates altogether in Croatian.

The new Serbo-Croatian standard created a situation where is Serbian works pre-1850 are unreadable to modern Serbs. Vuk's new standard essentially created a new language for Serbs that has no historical continuity or basis. Nevertheless, the longstanding tradition "writing the way you speak" continues to this day in Serbia despite the presence of a standard language.

The modern Serbian standard is a combination of the East Herzegovinan Ijekavian standard from 1850 combined with the official Serbian language spoken by the Serbian elite which is based on the

Ekavian Štokavian dialects of Vojvodina-Sumadija in the far north of Serbia. Ekavian was originally spoken only in Vojvodina-Sumadija, but it has now spread through Serbia to become the most common dialect. Ijekavian Štokavian tends to be spoken as a mother tongue more by Serbs in the west and in Southeastern Croatia. There are also some Ikavian Štokavian speakers in the far north of Vojvodina and across the border into Hungary.

Although the official standard Serbian exists, the standard is not enforced, and most people continue to use their mother tongue, usually Ekavian Štokavian.

Both Neo-Štokavian and Old Štokavian are spoken by Serbs, but Serbs overwhelmingly speak Neo-Štokavian, and Old Štokavian is only spoken in a few places Slovenia and even fewer places in Serbia.

Serbian Old Štokavian has three branches in Serbia, Bela Krajina Old Štokavian, Kosovo–Resava Old Štokavian and Torlakian Old Štokavian.

Kosovo–Resava Serbian or Older Ekavian is an Old Štokavian dialect spoken by Serbs in and around Kosovo, in Greater Moravia in Southwestern Serbia, in Northeastern Serbia around Smederovo, and in part of the Banat in Serbia and Romania.

Torlakian or Timok-Prizren Serbian or Old Serbian is an Old Štokavian dialect that is spoken in the south of Serbia and is probably a separate language. It is dealt with in a special section below.

Bela Krajina Old Štokavian or Bela Krajina Serbian is an Old Štokavian Ijekavian-Ekavian dialect spoken in Bela Krajina near Metlika in the far south of Slovenia in Paunoviči, Marindol, Miliči, and Bojanci by Serbs who migrated there long ago. These people are called Usoks. This dialect has come under serious Slovene influence. They continued to speak Serbo-Croatian and historically did not intermarry with their neighbors. While Slovenia is almost completely Catholic, these Serbs are Eastern Orthodox.

These people came up from the south during the Turkish Wars. Another group from this same conglomeration went to the Zumberak Hills. The Serbs in the Zumberak Hills were Croatized because these people now are Catholics and Greek Catholics (M. Jembrih 2019).

In Medieval times, Bela Krajina was in the Kaikavian zone (M. Jembrih 2019). They even [started to call themselves *Horvati*](#), after *Horvatski*, the name of Kaikavian until 1850 when Zagreb proclaimed Neo-

Štokavian as the official language of Croatia (Golec 2012).

The dialect seems to be dying out. It is mostly spoken in Marindol, Miliči and Paunoviči, while in Bojanci the oldest generation are monolinguals, the middle generation speaks Bela Krajina and Slovene equally, and the youth have only passive competence in the dialect. Their elders speak to them in Bela Krajina, but the young people answer back in Slovene.

No information on MI with Bela Krajina Old Štokavian and the rest of Štokavian is available. However, the majority of the people in Bela Krajina are Catholic Kaikavian speakers living in Slovenia (M. Jembrih 2019). A lot of them were originally Serbian Usoks but they became Croatianized via conversion to Catholicism or Greek Catholicism.

Gornji Milanovac Serbian appears to be the same dialect as Bela Krajina. It is spoken in Gornji Milanovac near Čačak in the Moravica District in Central Serbia in the West Morava Valley in the Šumadija region. This may have been where the Usoks originated from before they fled north to escape the Turkish attacks and moved north to Slovenia in 1528-1593.

Serbian Neo-Štokavian has three branches in Serbia, Ekavian Štokavian, Ikavian Štokavian, and Ijekavian Štokavian.

Ekavian Štokavian is best seen as part of a single dialect – perhaps called *Continental Štokavian* (Kolanović 2015). There is one branch of Ekavian Štokavian in Serbia, *Vojvodina-Sumadija Ekavian Štokavian* or Younger Ekavian.

Serbian Štokavian, Serbian Ekavian Štokavian, or Younger Ekavian is a Neo-Štokavian dialect originally spoken only in the Vojvodina-Sumadija region as the Šumadija-Vojvodina dialect, but it has now spread through the country to become the most popular dialect, as it is the dialect of the Serbian elite, and is now spoken by 60% of the country. It is also spoken by Serbs in the west and northeast of Kosovo and on the Danube River in Banatska Klisura (Clisura Dunării) in Romania.

Šumadija-Vojvodina Serbian is the Serbian Ekavian Štokavian base for the Standard Serbian Ekavian variant. It is spoken across Vojvodina and around Šumadija in Northern Serbia, in Northwestern Serbia, in Belgrade and in a few towns in Croatia around Vukovar.

Smederevo-Vršac Serbian is the main subdialect of Šumadija-

Vojvodina Serbian spoken in Serbia. It is spoken in Central Serbia from Belgrade to the eastern border.

Novi Sad Serbian is a dialect of Šumadija–Vojvodina Serbian spoken around Novi Sad in Northwestern Serbia.

Vojvodinan Serbian is a dialect of Šumadija–Vojvodina Serbian spoken in Vojvodina in far northeastern Serbia.

Serbian Ikavian Štokavian is a Neo-Štokavian dialect that is much less common than Ekavian Štokavian in Serbia, being spoken in only a few places. Speakers are called Bunjevacs, and they are best seen as part of a separate language called Bunjevaci Ikavian Štokavian or Bunjevaci, part of a larger language called Ikavian Štokavian. They speak New Ikavian or Younger Ikavian.

Serbian Ijekavian Štokavian is a Neo-Štokavian dialect spoken in three branches in Serbia: *Užice-Zlatibor Ijekavian Štokavian*, *Western Ijekavian Štokavian*, and *East Herzegovinan Ijekavian Štokavian*. One branch, *East Herzegovinan Ijekavian Štokavian*, is spoken as one of the two standard languages, the other being a standardized Serbian Ekavian Štokavian. Another, *Užice-Zlatibor Ijekavian Štokavian*, is spoken in the Sanjak. The third, *Western Ijekavian Štokavian*, is spoken in a few places in Western Serbia near Croatia.

Užice-Zlatibor Ijekavian Štokavian, *Užice-Zlatibor Serbian*, *Užican*, or *Zlatiborian Serbian* is an Ijekavian Štokavian dialect spoken in Southwestern Serbia in an area called the Sanjak by a group called the Sanjaks, many of whom are Muslim. Many people imply that it is fully intelligible with Standard Serbian, but a few people imply there might be difficult intelligibility with the harder forms spoken by rural peasants. For now, it is best to leave it as a Serbian dialect.

Western Ijekavian Štokavian is spoken in a few parts of Western Serbia near the Croatian border.

East Herzegovinan Ijekavian Štokavian is the basis for Standard Serbian. It is spoken throughout the country, but it is less popular than Younger Ekavian, the other standard language.

Serbian intelligibility of other Slavic lects:

Serbian intelligibility figures benefited from the presence of many good, objective Serbian informants and a lack of nationalist contamination.

Serbian has superb intelligibility of all of Serbo-Croatian for some reason. For that reason, we might consider it to be the most "central" Serbo-Croatian dialect of the Standard Serbo-Croatian language.

Serbian has 99% intelligibility of Bosnian and Montenegrin.

Serbian has 97% intelligibility of Croatian.

Serbian has 84% of Virje-Djurdjevac Kaikavian, a Kaikavian dialect transitional to Croatian.

Serbian has 77% intelligibility of Torlakian.

Serbs have 72% of Kaikavian. This figure is remarkably high, but that is due to the restricted number of reports. I only had two Serbian-Kaikavian intelligibility reports, and in each of them, Serbs were evaluating Kaikavian-Štokavian transitional dialects that have very high MI with Serbo-Croatians. If Serbs were presented with the hard Kaikavian dialects with poor MI with Serbo-Croatian, their score should drop a lot, probably down to 35-40%, where the Croats are.

Serbian has 57% understanding of Macedonian. It varies from 25% in Vojvodina in the north to 91% in Nis in the south, which is extremely variable range for a single small nation. This shows how dramatically different the Macedonian-Bulgarian-South Serbian ([Bulgarized](#)) system south of South Slavic from the Shtokavianized and Croatianized (Serbo-Croatianized) system in North Slavic.

Serbian has 55% of Pannonian Rusyn, which is basically Eastern Slovak.

Serbian has 30% of Bulgarian, quite low for two languages next to each other. Serbian has remarkably low intelligibility of Russian, but Croats score about the same here as the Serbs.

Serbian intelligibility of Russian is 12%, but Croats score about the same. Serbian has [low intelligibility](#) of Russian. The [vowel reduction](#) in Russian causes most of the intelligibility problems

Serbian has very low intelligibility Čakavian, 10%.

Serbian intelligibility of Ukrainian is remarkably low at 7%.The large amount of [Polish vocabulary](#) is what hinders intelligibility with Ukrainian and makes it even harder for Serbs to understand than Russian, as Serbs understand Polish very poorly.

Serbian has no intelligibility at all of Bednjanski Kaikavian (1/500

words understood), but Croats hardly understand it better at 3%. Bednjanski Kaikavian must be one of the Slavic languages that is most poorly understood by other Slavic speakers, which makes sense if its substratum is the remains of the Indo-Iranian language spoken by the original Croats.

Serbs understand nothing at all of Rhodope Mountains Pomak, which also has a heavy dose of this pre-Slavic substratum, and it is as poorly understood by Slavic speakers as Bednjanski Kaikavian.

Torlakian

Torlakian is spoken in the south and southwest of Serbia, the far south of Kosovo, in Northern Macedonia, in far Eastern Bulgaria and in Eastern Romania and is transitional to either Macedonian, Serbian, or Bulgarian depending on the Torlakian dialect in question. Torlakian speakers generally do not see themselves as speaking a separate tongue. In each nation they simply feel that they are speaking a dialect of the national language. Torlakians say they are speaking a Serbian dialect in Serbia, a Macedonian dialect in Macedonia, and a Bulgarian dialect in Bulgaria.

Bulgarian Torlakian has difficult intelligibility with Standard Bulgarian, Macedonian Torlakian may have difficult intelligibility with Macedonian Proper, and at least the purer forms of Serbian Torlakian are not intelligible with Serbian.

This conclusion will be quite controversial, but at the moment, there seems to be good evidence that Torlakian is indeed a separate language.

The Serbian vocabulary in both Macedonian and Torlakian is very similar, stemming from the political changes of 1912; whereas these words have changed more in Bulgarian.

Romanian Torlakian is spoken by a group of speakers in Eastern Romania around the town of Carasova by a group of people called Carasovans. Unlike most Serbs, Carasovans are Catholic. They speak an old Serbian Torlakian dialect from the [Timok Valley around Zaječar](#).

Macedonian Torlakian is traditionally spoken mostly by the Albanian

minority but also by a few Serbs in a band across Northern Macedonia around Tetovo, north of Skopje, Sveti Nicole, Kratovo, and Kumanovo. This group is said to speak either Bulgarian or Macedonian, but it is better to see them as speaking Torlakian. This dialect is best seen as closer to Macedonian than to Bulgarian or Serbian. Even Macedonians refer to it as [Macedonian transitional to Serbian](#), which is a good partial description of Torlakian.

Speakers of *Central Pelagonian Macedonian* on the Pelagonian Plain say that *Kumanovski Torlakian* spoken in Kumanovo in far Northwestern Macedonia is unintelligible. Macedonians say that Macedonian Torlakians speak something that looks more like Serbian. Found samples of Kumanovski speech were humorous sayings about how no one can understand them. And Macedonians say that Kumanovo speech is "unique and separate." This implies that they think it is not a part of their Macedonian language.

At the same time though, Macedonians will state that all of these Torlakian dialects are actually Macedonian. When Macedonians hear Serbian Torlakian speech, they think it is a [comedy routine](#) where someone is mixing up Serbian and Bulgarian as a joke.

Kosovan Torlakian was formerly spoken in a wide area of Western Kosovo before the war in 1999, but is now spoken only in the far south by Serbs only around Prizren on the Macedonian border and to the southwest in the area where Kosovo, Macedonia, and Albania all come together by Serbian Muslims called Gorani. The Serbs fled or were ethnically cleansed during and following the war.

[Serbian Torlakian](#), the Torlakian spoken in the southeast of Serbia and by the Serbian minority in Northern Kosovo is different. Serbian Torlakian is neither Bulgarian nor Serbian, instead it is a mixed Bulgarian-Serbian transitional dialect. The borders of Serbian Torlakian in Serbia run to near Southwestern Kosovo, to the northern border of Macedonia, to most of the northwestern border of Bulgaria.

Traditionally Serbian Torlakian is divided into three dialects, *Nis* (*Nislijski Torlakian*) in the north, *Leskovac* (*Leskovacki Torlakian*) in the center, and *Vranje* (*Vranjanski Torlakian*) in the south. Serbs in the north of Serbia around Belgrade, Novi Sad, and Smederovo state that Torlakians do not even speak Serbian; instead, Northern Serbs say that Torlakians speaks Macedonian.

Pirot Valley Torlakian is generally excluded from this classification, as the other Serbian Torlakian speakers say that Pirot Valley speaks Bugarashki (Bulgarian-like). Nevertheless, Pirot Valley is indeed part of the Torlakian linguistic system. This speech uses many Serbian borrowings for words that Torlakian lacks. Speakers in the Pirot Valley say they do not speak either Bulgarian or Serbian. Bulgarian and Serbian speakers report that Pirot Valley is hard to understand because of the presence of so many Romance words which came in via Romanized Thracian.

The difference between Vranje and *Belgrade Serbian* may be greater than the differences between all three forms of Croatian – Štokavian, Čakavian, and Kaikavian. Serbian is probably closer to Slovene than it is to Torlakian. If this is true, and there is a good chance that it is, then Torlakian is surely a separate language from Serbian, as Čakavian and Kaikavian are widely regarded by Croatian linguists as being separate languages.

Nevertheless, *Belgrade Serbian*, *Smederovo Serbian*, *Kosovo-Resavian Serbian*, Macedonian, Torlakian, and even Bulgarian are closer to each other than any of them are to Slovene or Kaikavian. This is because they all underwent Balkanization, which reached 2/3 of the way north in Serbia all the way up to Belgrade but not further north. In addition, Standard Serbian itself is Balkanized.

However, one can make a good case that Serbian dialects in the very far north of Serbia (*Vojvodina Serbian*) are closer to Kaikavian than to Torlakian.

Nislijski Torlakian is in poor shape in recent years and has been undergoing a period of Štokavization for a long time now.

This has taken some interesting forms. While Nislijski lost morphological or flexional cases in line with the Bulgarian and Macedonian loss of cases, the strong influence of the standard language resulted in the lost cases actually being reintroduced. While case is still going strong in Serbian, even there the trend there also is towards loss of case. Locative and Dative cases have merged so completely that some Serbian grammar books decline to distinguish them. Nominative and Accusative cases for masculine and neuter nouns have collapsed together.

The loss of case in Southern Slavic actually happened over centuries.

It began among speakers of Slavic dialects of Northern Greece where these speakers came under the influence of a general Balkan Sprachbund. It then spread north to Macedonia and Bulgaria and as far as Southern Serbia. The spread further north was prevented by the institutionalization of a Serbian state with Štokavian chosen as the official language. Although case is returning to southern Serbia, it will not be reinstated in Bulgaria and Macedonia since those lands instituted formal states with official languages that will prevent the spread of case reintroduction.

Nislijski is now a mixture of traditional Nis speech with heavy Štokavian influence. The old Nislijski is only known from some poems, songs, sayings, etc., some of which the speakers themselves do not even understand. The fact that modern Štokavized Nislijski speakers often cannot understand traditional Nislijski seems to imply that even *Northern Torlakian* has difficult intelligibility with Serbian.

There are groups of Bulgarians living in Serbia in the areas of Basilograd and Dimotrovograd who speak what people call a Bulgarian-Serbo-Croatian transitional dialect, and Serbs are able to understand these Bulgarians well. These Bulgarians are apparently speaking Torlakian. These Bulgarians have a weak Bulgarian identity and they only study Bulgarian in Grades 1-4 for 3-4 hours per week.

Bulgarian, Shop, or Tran Torlakian is another major Torlakian dialect which spoken in the northwest of Bulgaria around Belogradochik, Godevo and Tran, and a dialect close to Torlakian is even spoken in the capital Sofia. This dialect is also called *Shop Bulgarian*. Shop Bulgarian is the same thing as *Tran Bulgarian*. Formally, both are classified as Bulgarian dialects. 90% of the speakers refer to themselves as Macedonians, but this is a very emotional subject with Bulgarians, who deny that they identify themselves this way.

Speech to the north in the far northwest corner of Bulgaria is called *Vidin-Lom Bulgarian*, a Bulgarian dialect. Although linguists do not put this dialect inside formal Torlakian, perhaps it is better placed there. Bulgarian speakers say it is hard to understand.

To the southwest of Sofia, *Samokovo Southwest Bulgarian*, classified as a Bulgarian dialect, is spoken. Standard Bulgarian speakers say it is also hard to understand.

Although linguists put Samokovo Southwest Bulgarian outside of

Torlakian and inside of Bulgarian dialects instead, these dialects are both better seen as Torlakian.

Shop ranges from the Rila Mountains north to around Sofia all the way up to Vidin on the Danube.

This, like the Prizren-Timok Torlakian in Serbia, is characterized by mostly Bulgarian phonology and mostly Serbian vocabulary. Bulgarian speakers in cities such as Varna and Bourgas refuse to see this lect as Bulgarian; instead they say that these people speak Serbian.

Bulgarian speech in Sofia is also disliked in the rest of the country which sees even the speech of the capital as Macedonian/Serbian. Sofia no longer speaks Shop, but many surrounding villages do. Sofia now speaks Standard Bulgarian with some Shop characteristics.

Other Bulgarians say that Bulgarian Torlakian is barely intelligible to them. Part of the problem is that Standard Bulgarian is half-based on Eastern Bulgarian dialects, further distancing the standard from Bulgarian Torlakian. The standard was based in part on the *Veliko-Turnovo dialect* because this city was the capital of Bulgaria during the pre-Ottoman Slavic Kingdom of Bulgaria.

Torlakian is a heterogeneous language. One group in Macedonia and Serbia speaks a dialect that is closer to Macedonian than to Serbian or Bulgarian, and another in Bulgaria seems to speak more of a Bulgarian-Serbian mixed dialect. The speech of the Pirot Valley in Serbia has difficult intelligibility with both Serbian and Bulgarian.

All in all, the various forms of Torlakian in Kosovo, Macedonia, Serbia, and Bulgaria seem to be more part of Bulgarian-Macedonian than part of Serbo-Croatian (Moseley 2010), and they seem closer to Macedonian than to Bulgarian.

Nevertheless, speakers of Standard Macedonian say that even Kumanovo Torlakians who speak the Torlakian form closest to Macedonian are unintelligible.

On the other hand though, the MI evidence for Macedonian-Torlakian is controversial. Macedonian informants report 93% intelligibility of Serbian Torlakian and 95% intelligibility of Bulgarian Torlakian. Serbian Torlakians report 92% intelligibility of Macedonian. This implies, as above, that Torlakian has an especially close relationship with Macedonian or possibly that Torlakian is simply a dialect of Macedonian spoken in Macedonia, Bulgaria, and Serbia.

On the sociolinguistic end, Croatians say that they understand Torlakian (60%) much better than they understand Macedonian (25%) (Petričević 2019). This wide discrepancy in intelligibility implies separate languages. This also implies that Torlakian is much closer to Serbo-Croatian than Macedonian is. Furthermore, Croatian informants state that when they listen to Macedonian and Torlakian, it does not sound like they are listening to the same language (Petričević 2019).

In addition, Macedonians have a much better understanding (77%) of Bulgarian than Serbian Torlakians do (40%). This wide discrepancy once again presupposes a separate language. In addition, it looks like Macedonian is much closer to Bulgarian than Torlakian is.

To sum up, Torlakian and Macedonian have a very close relationship, with Macedonian being much closer to Bulgarian and and Torlakian being much closer to Serbo-Croatian.

The existence of Torlakian as a separate language would be hard enough sell in this hyper-nationalistic region. On the other hand, the notion that Macedonian is spoken in large parts Bulgaria and Serbia would probably enrage Bulgarian nationalists.

The Bulgarians would be particularly upset because even Bulgarian linguists are staunch nationalists who insist that Macedonian is a Bulgarian dialect, despite the abundant evidence that this is not so. And the notion that this foreign language called Macedonian is spoken in their own land would be particularly insulting.

Serbians on the other hand seem much more intrigued by the notion of Torlakian as a unique Slavic lect. Serbians would have to accept that a Macedonian dialect is spoken in their land, but given the close relationship between Serbian and Macedonian nationalists, this might not be a problem.

Macedonians are furious nationalists, especially when it comes to Linguistics, and they seem to aggressively resist the very notion of a Torlakian language at all.

Looking at the picture from an intuitive lens, Macedonian Torlakian, Bulgarian Torlakian, and the pure Serbian Torlakian varieties of Southern Serbia do seem to be part of something we could reasonably call a Torlakian language. The heavily Štokavized Torlakian or Nislijski now spoken in Nis is no longer a part of the Torlakian language and is now a Serbian dialect.

Unfortunately, the status of Torlakian is highly politicized by Macedonian, Bulgarian, and Serbian nationalists, making objective research difficult.

For the time being it seems rational to set aside Torlakian as a unique and coherent Slavic lect, while the question of whether it is a separate language or a Macedonian dialect will have to remain up in the air.

Torlakian intelligibility of other Slavic lects:

Torlakian intelligibility calculation benefited from having Torlakian informants, including one excellent personal informant.

Torlakian has 90% inherent intelligibility of Macedonian. Of course this brings up the question of whether Torlakian is a dialect of Macedonian or not, a question that at the moment remains vexing and unresolved.

Torlakian has only 77% understanding of Serbian, especially the Vojvodina and Šumadija dialects in the north.

Torlakian has only 45% understanding of Serbo-Croatian, which probably means Croatian. Written understanding is also very poor at 40%.

Torlakian has 30% understanding of Bulgarian.

Torlakian has poor understanding of Slovene at 20%.

Torlakian understanding of Russian is also poor at 20%. Written intelligibility of Russian is a lot higher at 75%. This doesn't make much sense until you realize the closeness of Torlakian to Bulgarian, a language with a known connection to Russian, and to Serbian, a language which until 1850 had been half Old Church Slavonian. Of course OCS is a huge part of the Russian and Bulgarian languages.

Torlakian has very low understanding of Kaikavian at 15%, showing that it is completely outside of the Slovene-Kaikavian Northwestern South Slavic regional speech.

Torlakian has very low understanding of Czech at 12%.

Torlakian has the same understanding of Slovak at 12%.

Torlakian has extremely low intelligibility of Polish at 5%. It looks like the entirety of West Slavic is far beyond the speech of the Torlaks far in the South Slavic South.

Macedonian

Macedonian and Bulgarian would be much closer together except that in recent years, Macedonian has been heavily influenced by Serbian, and Bulgarian has been heavily influenced by Russian.

Indeed, Macedonian and Bulgarian used the same literary language up until the 1800's (though the local dialects no doubt differed from the standard). Scholars from Bulgaria and Macedonia met and decided what features should go into the literary language, making a lot of compromises. This went on until later in the century when the Macedonians experienced an upsurge in national awakening. Then Macedonian scholars decided that they did not want to speak the same literary language as Bulgarians. The two groups of scholars split, and the Macedonian literary language began to separate from Bulgarian.

70-100 years ago, Macedonian and Bulgarian were much closer together, and you could make a case that Macedonian was indeed a Bulgarian dialect. People who still speak the old Macedonian dialect from those times think the idea of a separate Macedonian language is absurd, and they resent Macedonia for trying to impose an artificial language on them.

The Macedonians who were expelled at gunpoint by Greek nationalists before World War 1 often still speak the old dialect. Greek troops arrived and ordered them to declare themselves Greek and speak Greek. People who did not speak Greek were often murdered on the spot. Others were given one day to pack their things and leave or die. 200,000 packed on a day's notice and marched out of the country on a modern day Trail of Tears.

While the Macedonian spoken in the north of Macedonia sounds very Serbian, Macedonian spoken in eastern Macedonia sounds quite Bulgarian, so Macedonian, like so many other Slavic languages, appears to be part of a dialect continuum.

In recent years, Macedonian has taken on quite a few Serbian borrowings, so it is moving away from Bulgarian.

There are many dialects of Macedonian, but in general the MI between them is full. The old dialect of the capital, Skopje, is dying out but it is still widely spoken all over the city. There is [controversy](#) about whether this dialect is downgraded or not. It [appears](#) that in very formal settings, it's best to use Standard Macedonian, but in any non-formal setting, it's acceptable to use Skopian.

Aegean Slavic dialects mostly spoken in Greece and to some extent in Albania such as the *Lerin-Kostur* and *Solun-Voden* dialects are the subject of ferocious debate regarding whether they are Bulgarian or Macedonian.

The discussion is very confusing and is made worse because much of the discussion has been deeply politicized. Nationalists on both sides claim all of these dialects for either Bulgarian or Macedonian.

The best analysis is that most Aegean Slavic dialects are Macedonian with only the *Ser-Drama* dialect in Greece and Bulgaria and the *Maleševo-Pirin* dialects in Greece, Macedonia, and Bulgaria being Bulgarian. The *Solun-Voden*, *Shtip-Strumica*, *Central*, *Lower Prespa*, and *Korča* dialects in Macedonia and Greece and the *Kostur*, *Upper Prespa-Ohrid*, *Vevčani-Radožda*, *Drimkol-Golo Brdo*, and *Debar* dialects in Albania are Macedonian.

Macedonian nationalists claim that the Maleševo-Pirin dialect is Macedonian, but some also [admit](#) that Bulgarian is spoken in parts of Macedonia.

Although Solun-Voden is a dialect of Macedonian, it sounds more like Standard Bulgarian than Standard Macedonian. This leads even Macedonians to say that the Solun-Voden speech around Bogdanci and Gevgelija [is Bulgarian and not Macedonian](#). This is not correct but at least they admit that Bulgarian is spoken in their country, which is correct.

Intelligibility data with either Macedonian or Bulgarian is not available on the Aegean dialects.

Kostur-Korča Macedonian of Northeastern Greece may have difficult intelligibility with Standard Macedonian, as it is considered to be the most diverse of the "Macedonian" dialects, apparently even more diverse than Kumanovo-Kratovo dialect, which Macedonians already regard as unintelligible. If Kostur-Korča is even further removed from the standard than Macedonian Torlakian, it may have difficult

intelligibility with the rest of Macedonian. For now will remain a Macedonian dialect. Bulgarian has 75% intelligibility of Kostur-Korča.

Macedonian intelligibility of other Slavic lects:

Macedonian intelligibility figures benefited from a good number of Macedonian informants, but nationalism caused some impairment, particularly on the hot-button Torlakian issue. A number of hostile nationalist informants had their judgments rejected. There were no good personal informants.

Although Macedonians claim difficult intelligibility of Macedonian Torlakian, they also claim good intelligibility of Serbian Torlakian and Bulgarian Torlakian at at least 91% and sometimes higher. This doesn't make much sense and adds to the mystery of what to do with Torlakian – whether it is a separate language or a part of Macedonian.

Macedonians have 70% intelligibility of Serbo-Croatian, and they can probably understand Serbian easier than they can understand Croatian. Macedonians have excellent full written intelligibility of Serbo-Croatian at over 90%.

The intelligibility of Macedonian with Serbo-Croatian is highly controversial and politicized. Most Macedonians already are able to speak Serbo-Croatian well. In fact, many Macedonians are switching away from the Macedonian language towards Serbo-Croatian. This gives rise to claims of Macedonians being able to understand Serbo-Croatian very well, however, much of this may be due to bilingual learning.

In fact, studying Macedonian intelligibility of Serbo-Croatian is difficult because 90% of Macedonians know at least some Serbo-Croatian. Even worse, almost 100% of young Macedonians can speak [perfect Serbo-Croatian](#). For instance, [on the radio in Macedonia](#), there are more Serbo-Croatian than Macedonian songs.

The Macedonian spoken near the Serbian border is heavily influenced by Serbian and is quite a bit different from the Macedonian spoken towards the center of Macedonia. One way to look at Macedonian is that it is a Serbian-Bulgarian transitional language. If even Macedonian-Serbian transitional Macedonian Torlakian is regarded by other Macedonian speakers as incomprehensible, then surely they understand pure Serbian even worse. Macedonians understand Serbo-Croatian [better](#) than they understand Bulgarian.

Macedonian intelligibility of Bulgarian is lower than the other way around, at 65%. Written intelligibility is a bit better at 75%. However, Macedonians can pick up Bulgarian easily, often after only a month of studying. Many Macedonians are attending Bulgarian universities.

Macedonians have 43% intelligibility of Czech. Written intelligibility is much higher at 79%.

Macedonian intelligibility of Slovak is 36%. Written is dramatically higher at 67%.

Macedonians have 27% intelligibility of Polish. Written understanding is somewhat better at 40%.

Macedonians have low intelligibility of Slovene at only 27%. Written understanding is much better at 73%.

Macedonian intelligibility of Russian is poor at 18%. However, they understand written Russian much better at 41%.

Macedonians intelligibility of Ukrainian is very low at 13%. Their understanding of written Ukrainian is much better at 27%.

Macedonian intelligibility of Kashubian is worst of all at 12%.

Bulgarian

Bulgarian is an odd example of a Slavic tongue with heavy Turkic influence. In the 7th Century when the a Turkic tribe under Asparukh invaded the area and gradually become assimilated and Slavified over time. This Slavified Turkic then mixed in with the language of the original Bulgarians who were already speaking an early Slavic language. The strong Turkic influence on Bulgarian means that the language is full of Turkic words. However, most of this vocabulary comes not from Asparukh's invaders in the 600's but from the centuries under which Bulgarian was under Ottoman rule.

Bulgarian along with Macedonian are also interesting in that they have lost most of their inflectional case and are now much more analytic than the rest of Slavic. The loss of case began in Aegean Slavic around Thessalonika in Greece and proceeded north forming the Balkan linguistic Sprachbund.

One theory is that the loss of case was caused by Greek influence.

Yet the truth is that no one knows what caused the features of the Balkan Sprachbund. Various theories have pointed to Paleobalkan languages, Greek, or Balkan Vulgar Latin. At the moment, none of these hypotheses are popular, but it is clear that this is contact-induced convergence that has its origins in massive bi- or multilingualism and reciprocal influences of various different languages on each other. The primary languages involved were Paleobalkan Slavic languages feeding back and forth into and out of each other (Strazny 2013, p. 118).

Bulgarian has a number of peculiarities that set it apart from other Slavic languages. While most Slavic tongues lack articles, in Bulgarian, the definite article is an enclitic suffixed to the noun.

There are different theories about where this enclitic came from.

One version is that it derived from Turkic influence, possibly from the Proto-Bulgaro-Turkic Bulgar language formerly spoken in this area before it became Slavified.

However, it could just as easily have come from Romanian. Romanian is the only Romance language with enclitic definite articles (Weigand 1925).

Bulgarian also has an irrealis-type mood known as nonwitnessed or inferential. This is either a fourth mood (Kutsarov 2007, pp. 282-286) or an "evidential," part of the inflectional system. Nonwitnessed has several forms known either as tenses or evidentials: renarrative, doubtful, inferential, and retold. Evidential is a way of inflecting the verb similar to aspect, tense, and mood.

The inferential forms are derived from Turkish (Gerdzиков 2003).

In addition, Bulgarian has retained the ancient verbal system from Proto-Slavic.

Bulgarian also lacks an infinitival verb form. The lack of an infinitive may also come from Romanian (Weigand 1925).

As in Poland, linguistics in Bulgaria has long been affected by nationalist discourse. Bulgarians are some of the worst ethnic nationalists in the Slavic region. For a long time, Bulgarian linguists have drawn maps of Bulgarian dialects covering not only the entire territory of Bulgaria but also across all of Macedonia and much of Northern Greece. These maps are based more on politics than reality.

In 1963, Asim Peco, the father of Bosnian linguistics, stated emphatically that Macedonian was clearly a separate language from Bulgarian.

The language question among Bulgaria and Macedonia has become very heated and there has been a lot of angry discourse between the nations. Inside Bulgaria, for a period from 1947-1956, Macedonians were acknowledged as an ethnic minority in Bulgaria. They were even given separate schools.

This had to do more with politics than anything else as the Communist leaders of Bulgaria and Yugoslavia tried to create some sort of a federation in the Balkans. This project is regarded as a catastrophe by Bulgarian nationalists, who are the vast majority of the population. This project was abandoned in 1956 and all rights of Macedonians were reversed. Macedonians were forced to identify as Bulgarians.

At one point a legal Macedonian party was formed. However, after the fall of Communism a new Constitution was written that outlawed all ethnic political parties, and a judge ruled that this Macedonian party was illegal. Since then, Macedonian activists in Bulgaria have been harassed, threatened and arrested for their activities. Popular opinion is wildly opposed to the idea that ethnic Macedonians exist in Bulgaria and almost everyone says the Macedonians are simply Bulgarians. Almost all Macedonians in Bulgaria have been identifying as Bulgarian in increasing numbers since World War 2.

It is certainly true that Macedonian is very intelligible to Bulgarians, more than the other way around for some reason. Most older Bulgarians can understand Macedonian perfectly, and most of the youth think it is just sounds strange and "old-timey", as in they use a lot of older words that have fallen out of modern speech. Macedonians are never translated on Bulgarian TV and only sometimes are they even subtitled.

In time though, the standard languages continue to drift further apart.

While reporting Bulgarian MI claims, an attempt was made to look at a great number of diverse sources. Not all Bulgarians are ethnic nationalists, and those that are not report much lower MI figures than those who are. Hence an attempt was made to discover which

responses were coming from Bulgarian ultranationalists. Once discovered, all judgments seen as coming from Bulgarian nationalists were discarded due to probable bias.

The *Ser-Drama-Lagadin-Nevrokop dialect* in Northeastern Greece and Southwestern Bulgaria and the *Maleševo-Pirin dialect* named after the Maleševo and Pirin Mountain Ranges in Eastern Macedonia and Western Bulgaria are transitional between Bulgarian and Macedonian.

The question of whether these dialects are Bulgarian or Macedonian has been deeply politicized and hence is hard to untangle.

Nevertheless, the best analysis with the most critical consensus seems to be that both are Bulgarian dialects transitional to Macedonian (Sussex and Cubberley 2011, p. 510).

Bulgarian is very dialectally diverse. In the far west of the country, so-called Bulgarian dialects are better seen as part of a language called Torlakian. Even outside of Torlakian, Bulgarian is heterogeneous. Two old men, one from Sofia and one from Burgas in the far east on the Black Sea Coast, will not be able to understand each other. However, the old man from Sofia may be speaking more Torlakian than Bulgarian. Whether this east to west miscommunication warrants further splits in Bulgarian is not known, but for the time being, it is best to only split off Torlakian.

Sofia Bulgarian, the dialect of the capital, Sofia, is stark in its difference. Although it is a prestige dialect, many Sofians think it is more prestigious and educated to speak Standard Bulgarian. It also has to do with education, and reading, especially literature. Bulgarian literature is written in Standard Bulgarian. The more one goes to school or reads Bulgarian literature, the more one has the tendency to speak the local dialect. People are moving to Sofia from all over the country, and they are bringing their local dialects with them. The result is a tendency to swamp out the local city dialect while at the same time promoting Standard Bulgarian.

However, Bulgaria is currently experiencing [dialect leveling](#) across the country. Dialect leveling is also happening in Macedonia, but the leveling in Bulgaria is [more severe](#). Part of the reason for this is television. All Bulgarian TV is in Standard Bulgarian. In Macedonia and Serbia, it is common for foreign productions to receive subtitles instead of dubbing, but all foreign production in Bulgaria is dubbed by

voice actors in Sofia who speak perfect Standard Bulgarian. Nevertheless, as in other Slavic countries like Croatia and Czechia, most Bulgarians use Standard Bulgarian with some dialectal features. Standard Bulgarian was a compromise composed of both Eastern and Western Bulgarian, which can be remarkably different.

Western Bulgarian intelligibility of other Slavic lects:

Western Bulgarian has 82% of Macedonian, a very high figure which explains why so many Bulgarians continue to insist that Macedonian is a Bulgarian dialect.

Western Bulgarian has 80% of Serbo-Croatian. This figure is very high, much higher than for Standard Bulgarian, and it shows the close relationship of Western Bulgarian with Serbian, the explanation of which is explained by the Torlakian language.

Eastern Bulgarian intelligibility of other Slavic lects:

Eastern Macedonian has only 35% intelligibility of Serbian, which shows how far it is from Western Bulgarian.

Eastern Bulgarian has only 30% intelligibility of Serbo-Croatian, dramatically lower than the figure for Western Bulgarian. Really it is only Western Bulgarian that has a close relationship with Serbo-Croatian.

Eastern Bulgarian has only 25% understanding of Croatian, significantly worse than for Serbian. This shows how far the Eastern South Slavic system is from the Western South Slavic system.

Eastern Bulgarian has only an incredibly low 12% comprehension of Western Macedonian. Looking at this figure alone, there is no way you can say that Bulgarian and Macedonian were the same language. Really what we are seeing here is the opposite ends of the Bulgarian-Macedonian dialect chain, the eastern and western ends of which can barely understand each other at all.

Bulgarian intelligibility of other Slavic lects:

Bulgarian intelligibility figures were good with many Bulgarian informants. There was only one personal informant, but unfortunately, he was only slightly useful.

Bulgarians can understand Macedonian to a great degree (82%) but not completely. A number of Bulgarians claim full understanding of

Macedonian. Eastern Bulgarians understand only a very low amount of Western Macedonian at 12%, a remarkably poor figure. However, the sample size was small. In contrast, Western Bulgarian has near-full of 82% intelligibility of Macedonian. Written is the same at 82%.

Bulgarians have a very high but not full understanding of Banat Bulgarian spoken in Romania at 80%.

Bulgarians understand Rup Bulgarian, a Bulgarian dialect, at a very high but not full level of 78%.

Bulgarian understanding of Kostur-Korča Macedonian is 75%. Kostur-Korča spoken far over by Albania and is one of the most divergent Macedonian dialects.

Bulgarians have 73% understanding of Zlatograd Rup Pomak, a dialect of the Pomak language. Bulgarian understanding of Pomak dialects is variable, ranging from high to near zero.

Bulgarians report partial understanding of Serbian at 60%. This is much higher than what scientific testing found for Serbo-Croatian, but the sample size was small. Nevertheless, Southern Serbian is a lot like Bulgarian. Eastern Bulgarian understands only 35% of Serbian, which is closer to what the scientific finding found. This can be explained because Standard Bulgarian is based on Eastern Bulgarian.

Bulgarians also have partial comprehension of Croatian at 50%. But Eastern Bulgarian has only 25% understanding of Croatian, about what scientific testing found. This makes sense because Standard Bulgarian is based on Eastern Bulgarian. This is much higher than the figure that scientific testing arrived at for Serbo-Croatian and the figure is hard to explain. However, the sample size was very small.

Bulgarian intelligibility is 46% of Russian. Written is much higher at 71%. In some cases understanding of spoken Russian can be very high, and some scholars put the figure at 75% (Karamian 2014). But the rate is highly variable, and some Bulgarians understand very little Russian, maybe 5-10%. Still, most Bulgarians can understand Russian well. Averaging together the various estimates, we get a figure of 46% intelligibility. The highly variable rate of 7-85% demands some sort of an explanation and implies there is some bilingual learning going on.

Nevertheless, Bulgarian-Russian intelligibility seems much exaggerated. Some Russians and Bulgarians say they understand

almost nothing of the other language. Most Bulgarians over the age of 30-35 understand Russian well since studying Russian was mandatory under Communism. This may well complicate the intelligibility figures and introduce a level of bilingual learning into the findings.

The closeness of Bulgarian and Russian came about because Bulgarian is based on OCS, and Russian is based on Old East Slavic. These languages were close to each other and could be seen as cousins. Due to its OCS origins, Bulgarian has retained much of its OCS vocabulary over the years. Hence even though the Bulgarian and Russian grammars are quite different, much of the vocabulary will be familiar to speakers of both languages.

The story of OCS itself is interesting. OCS itself was based on the Common Slavic dialect spoken in Thessalonika, Greece by Aegean Slavs in the 800's. Greek priests Cyril and Methodius translated the Bible from Greek into Common Slavic. They used words that were common to all of the Slavic lects in the area and added many Greek, Latin and Hebrew words (Strazny 2013, p. 117).

This OCS then became the basis for the literary Bulgarian language. Later this Bulgarianized version of OCS had massive input on the Russian language. Around 1700, Russian scholars decided that Russian had far too many Greek and Latin borrowings from the West. They wanted a more authentically Eastern tongue. They used this Bulgarian version of OCS to coin many new Russian words.

Bulgarian-Russian written intelligibility is much higher. Bulgarian and Russian written languages are close because the Ottoman rulers of Bulgaria would not allow printing in Bulgaria. Hence, many religious books were imported from Russia, and these books influenced Bulgarian. Russian influence only ended in 1878 (Karamian 2014).

Bulgarian intelligibility of Tran Torlakian, formally a Bulgarian dialect, is quite poor at 40%, but the sample size was very small. There are no figures for Macedonian or Serbian Torlakian

Bulgarians understand 35% of Ukrainian when spoken and 80% when written. The relatively high figure may be due to Ukrainian's similarity with Russian.

Bulgarian comprehension of Belarussian is the same as Ukrainian, 35%. The two languages are quite similar.

Bulgarian intelligibility of Serbo-Croatian is only 27%* (Golubovic and

Gooskens 2015a). Bulgarian understanding of written Serbo-Croatian about the same as oral, 32%* (Golubovic and Gooskens 2015a). Spoken intelligibility is 30% in Eastern Bulgaria, similar to above, and it is as high as 80% in Western Bulgaria. Western Bulgarian may resemble Torlakian, hence the high figure similar to Serbian Torlakian. The sample size from Eastern Bulgaria was small. That Eastern Bulgarian is lining up with the figure for Standard Bulgarian makes sense because Standard Bulgarian is based on an Eastern Bulgarian dialect. Part of the problem here is that Bulgarian is not spoken the same way it is written like the more phonetic written Serbo-Croatian is.

Bulgarians only understand 20% of Slovene* (Golubovic and Gooskens 2015a). Written understanding is the same at 21%* (Golubovic and Gooskens 2015a).

Bulgarians have a poor understanding of Čakavian at 20%. As in other cases, this makes sense if we can see Northern South Slavic and Southern South Slavic as completely different systems.

Bulgarian intelligibility of spoken Slovak is very low, 13%* (Golubovic and Gooskens 2015a). Written is somewhat higher at 25%* (Golubovic and Gooskens 2015a).

Bulgarian intelligibility of Czech (Golubovic and Gooskens 2015a) is 10%*. Written Czech is also poorly understood with a figure of 21%* (Golubovic and Gooskens 2015a)

Bulgarian understanding of Polish is very low, 10%*(Golubovic and Gooskens 2015a). Written is also low at 18% for Polish* (Golubovic and Gooskens 2015a).

Bulgarian intelligibility of Upper Sorbian is very poor at 10%.

It is interesting that Czech, Polish, and Upper Sorbian all come in at 10% understanding by Bulgarian. Both Czech and Polish are West Slavic, the furthest away from Bulgarian, and Upper Sorbian is between Polish and Czech. West Slavic Slovak is nearby at 13%.

Bulgarians comprehend Greek Pomak, a dialect of the Pomak language, at a very low level, 10%. As mentioned, Bulgarian understanding of Pomak dialects ranges to quite high to almost nothing.

Bulgarians understand Kashubian nearly the least of all at 6%.

Bulgarian intelligibility of Rhodope Mountains Pomak dialect of the Pomak language is extremely low at 3%, even though it is said to be a Bulgarian dialect. However, the sample size was very small. This is one of the most archaic languages in South Slavic, Archaic Islander Čakavian. Both are reminiscent of Early South Slavic. Perhaps Rhodope Mountains Pomak is Early Southern South Slavic.

Banat Bulgarian

Banat Bulgarian is an outlying Bulgarian dialect spoken in the Banat region which stretches from of Vojvodina in Serbia to Romania. There are only an estimated 3,000 speakers left in Serbia, but there are still 12,000 in Romania. In Serbia, they reside in villages with Hungarians and Slovaks.

Despite the time depth between the two dialects, the differences between Banat Bulgarian and Standard Bulgarian are not large.

The differences include more palatalization of sounds, including palatalization of final consonants not typical in the standard language, the plural suffix in *-e* rather than *-i*, a somewhat Romanianized phonology, and some archaic words. It also has quite a few German and Hungarian words. One notably thing about Banat Bulgarian is the language is written in the Latin alphabet whereas Standard Bulgarian is written in Cyrillic. Interestingly, most Banats are Catholics whereas Bulgarians are mostly Orthodox Christians.

The Banats were Bulgarian Catholics called Paulicians who fled their villages in Northern Bulgaria due to tragic events occurring in the Danube region of Bulgaria in the area of the towns of Svishtov and Nikopol in the 1600's and 1700's. At the end of the 1600's, a Catholic revolt called the Chiprovsky Rebellion broke out against the Ottomans.

Following this failed revolt and after a series of Ottoman-Austrian and Ottoman-Hungarian wars wracked the region, this group of Catholics fled the area around 1750 seeking a better life outside of the Ottoman Empire. After much persecution, they were able to negotiate a deal with the Austrian rulers Leopold, Carl VI, and Maria Theresia which

would allow them to settle in the regions of Transylvania and Banat. They were able to acquire land title after working land for thirty years. Their first years were beset by multiple natural disasters such as floods and suffered much hunger and poverty. After the Russo-Turkish War in 1887-1878 and the founding of the Principality of Bulgaria, some Banat Catholics returned to Bulgaria, this time settling in the West, especially in Bardanski Geran. But over the next 50 years, more catastrophes occurred, particularly World War 1, which caused many of this group to once again flee back to Banat.

Although at first glance intelligibility between Banat Bulgarian and Standard Bulgarian seems like it would be quite good, Bulgarians say that Banat Bulgarians do not speak their language. A comparison of the same text in Bulgarian and Banat Bulgarian shows that the two cannot possibly be mutually intelligible. There may be 80% MI between Banat and the rest of Bulgarian. Nevertheless, Banat is still close to Eastern Bulgarian dialects in general.

In 1994 an important Bulgarian dialectal study was done on the Bulgarian Catholic speech of the northern Bulgarian villages of Oresh, Trunchovitsa, Belene, and Malchika. The paper stated that this was the same speech as the Banat Bulgarians who had left several hundred years ago, but this was the group that stayed behind in Bulgaria. These dialects are no doubt closely related to and may be intelligible with Banat Bulgarian.

Banat Bulgarian is still learned by children in the Banat of Serbia. The group is very religious and all of their Catholic religious material is written in Banat Bulgarian with a special Latin script that covers all the sounds of the language. The children grow up reading the religious books of their mother tongue. Some who wish to immigrate to Bulgaria for study or work study Standard Bulgarian, classes of which are offered in the community. The fact that Banat Bulgarians need to study Standard Bulgarian before they move to Bulgaria implies that Banat Bulgarian is a separate language.

There are also speakers of Standard Bulgarian living in Serbia. Although they use Standard Bulgarian, they write in a sort of transitional dialect between the Standard Bulgarian and Serbian languages.

Splitting off Banat Bulgarian as a separate language would be

extreme, but Macedonian has already been split off and Torlakian is next if it is a separate tongue. Bulgarian linguists are all nationalists, and they are some of the worst nationalistic linguists of them all. Surely, splitting off Banat would be an outrage to Bulgarian nationalist linguists. But the task of linguists, even in the hot seat of the Balkans, is one of science and not politics.

Anyway it is quite clear that there are more languages than this under the Bulgarian umbrella. For instance Rhodope Mountains Pomak is not only a separate language from Bulgarian, but it is one of the most divergent Slavic languages of all. Bulgarian nationalists fight tooth and nail against recognition of Pomak as a separate language, but it has a better claim than Torlakian or even Macedonian. Anyway, science must always prevail over politics in the sciences.

West Slavic

The West Slavic languages share certain things in common. All of these peoples joined the Roman Catholic Church, hence their religious literature was written in Latin for a long time. Consequently a lot of Latin and Romance vocabulary entered these languages. The presence of so many Western words is one of the main things that differentiates these languages from East Slavic.

Slovak

Slovak is a dialect of the Czechoslovak language. It consists of two lects, Slovak and Eastern Slovak.

Although it is an extreme statement, this paper will take the point of view that Slovak is not in fact a separate language but instead is a part of a single language called Czechoslovak. This paper lumps or splits strictly on MI grounds and does not deal with political or sociological reasons for languagehood. According to MI, there is no linguistic reason to say that Slovak is a separate as it has 94% intelligibility of Czech proven by formal study (Golubovic and Gooskens 2015a). However, Eastern Slovak is surely a separate

language.

However, Slovak is also a macrolanguage consisting of Slovak and Eastern Slovak.

Slovak is one of the most conservative Slavic languages, although this is somewhat controversial. Slovak and OCS case endings look very similar.

As there was no center of Slovak culture until recently, Slovak dialects began diverging early both from each other and from their Slavic roots.

There was an attempt to create a common language in the late 1800's by using one of the dialects, but the dialects were so far apart that this was rejected by speakers of other dialects. A compromise was made, and as in Slovenia, an artificial official Slovak language was created based on all Slovak dialects but centered around Central Slovak. Instead of being based in part on literary languages as in Croatia, Slovenia and Czechia, the Slovak literary language was based completely on spoken speech (Rusinko 2003).

To make the standard, any unusual vocabulary and grammatical features were stripped from the language along with most of the irregularities and the Slovak phonology was rejected and Czech's was copied instead. The result was a much simpler and more regular language as compared to the spoken dialects.

The Austro-Hungarian Empire was bad for many of the unofficial languages within the imperium. As Czech was nearly driven extinct by German, so Slovak was nearly driven extinct by Hungarian.

Slovak has high dialectal diversity especially in contrast to Polish, possibly due to the very mountainous terrain. Dialects are still very strong in Slovakia (Kaplan and Baldauf 2005).

Slovak dialects are quite far apart. The jump between *Trenčín Western Slovak* and *Zvolen Central Slovak* is not far in geographical terms, but it is [equal to the difference](#) between *Trenčín Western Slovak* and *Moravian Eastern Czech*, between which there is a large gap.

[Western Slovak](#) consists of the *Kysuce*, *Trenčín*, *Trnava*, *Nitra*, *Bratislava*, *Zhilina*, *Chadka*, and *Záhorie* dialects. Western Slovak has so many Czech words that this has led to the mistaken conclusion

that Western Slovak and Czech are nearly the same language. They do appear close though. The absence of Czech ř is the most obvious difference (Zušťák 2015).

Western Slovak speakers use a lot of “Czechisms” or what they call “Bohemianisms” in their speech, often without realizing it. These Czechisms are often a Czech borrowing that has been changed by Slovak phonology. These Bohemianisms appear more in Western Slovak than in the rest of Slovak. People are often unaware that they are using a Czech borrowing that has been changed, and they do not know the original Czech word it was derived from (Zušťák 2015).

The main differences between Western Slovak and Czech are the absence of ě, ř, and ů - Slovak people usually cannot pronounce ř, and when they do, it is pronounced as ž - and the presence of ô and ä. For example, “meat”: Slovak *mäso*, Czech *maso*; “around”: Slovak *vôkol*, Czech *vůkol* (Zušťák 2015).

Compare for instance English “What it is good for?”: Czech *K čemu je to dobré?*, Bohemianized Slovak *K čomu je to dobré?*, Standard Slovak *Na čo je to dobré?* English “far more”: Czech *Daleko víc*, (Czech), Bohemianized Slovak *Ďaleko viac*, Standard Slovak *Omnoho viac* or *Oveľa viac* (Zušťák 2015).

[Moravian Slovak](#) is a Slovak dialect that is actually spoken in Czechia (Machek 1997), but most sources do not note this. In far Eastern Czech Republic, Moravian is spoken. Moravian is actually two dialects – Moravian Slovak, spoken in Czechia near the Slovak border (Machek 1997), and Moravian Czech, spoken in Czechia further to the west. In fact, the Moravian Slovak-speaking part of Czechia is often called *Slovácko*. The two dialects are very similar and highly intelligible with each other (Zušťák 2015).

Southwestern Slovak consisting of the *Trnava*, *Bratislava*, and *Záhorie* dialects, is nearly the same language as spoken in the Eastern Czech Republic. However, the phonology has been affected by Hungarian, since many Hungarians reside in this part of Slovakia (Zušťák 2015).

[Záhorie Slovak](#) is a Southwestern Slovak dialect spoken by people in the Záhorie Region in the extreme west of Slovakia in the Little Carpathian Mountains that is very close to Moravian Slovak across the border. It has many Czech borrowings that have been changed so

much that they do not resemble the original Czech word much. While Záhorie is very different from surrounding West Slovak dialects, intelligibility is full (Zušťák 2015).

Southern Slovak speakers on the Hungarian border have a harder time understanding Polish because speakers do not hear it much. Their intelligibility of Polish is ~10%. This implies that the higher Slovak-Polish intelligibility figures reported may be due to bilingual learning on the part of Slovaks, and Slovak-Polish inherent intelligibility is quite low.

Central Slovak, consisting of the Liptov, Martin, Prievidza, Levice, Luchenec, Banská Bystrica, Orava, Turiec, Tekov, Hont, Novohrad, Gemer, and Zvolen dialects, is the basis for Standard Slovak, which in turn was based on something called Cultural Standard Slovak, which had been spoken by townspeople and minor nobility since the 1100's. The first printed text in Cultural Standard Slovak was written in 1625.

In the 1800's, there were complaints that Protestant and Catholic Slovaks spoke such different dialects that they could no longer understand each other.

At the time, Slovak Protestants wrote in Czech, not in Slovak (Kaplan and Baldauf 2005).

Therefore, Standard Slovak was created by L'. Shtur and colleagues, an artificial language consisting of 50% Protestant Slovak speech and 50% Catholic Slovak speech. Central Slovak has full intelligibility of Czech.

Central Slovak is considered to be a sort of an Esperanto of Slovak and also of Slavic in general because it has so much in common with the rest of Slovak. Another reason for this may be its essentially "central" nature as a sort of center of gravity of Slavic. Some think that this notion of Central Slovak as an Esperanto of Slavic is a myth, but it has been endorsed by prominent Slavists such as Jernej Kopitar, František Kopečný, Karel Horálek, and Mario Pei.

However, other studies call this into question and suggest that Slovak is no better understood by non-Slovaks than any other Slavic language (Golubovic 2015)).

Central Slovak is one of the most interesting Slavic dialects because although it is formally placed in West Slavic, the general agreement is that it is a South Slavic language with a large number of West Slavic

features (Popović 1960).

The strong links between Central Slovak and South Slavic, in particular Kaikavian, have been noted by scholars (Nuorluoto 2010).

That Central Slovak is placed in West Slavic shows that South, West, and East Slavic groupings are based more on geography than linguistics. But this is true in many linguistic families – consider the “Siberian” grouping of Turkic, more of a geographical grouping than a linguistic one.

Central Slovak is also considered one of the most conservative Slavic dialects along with Eastern Serbo-Croatian (Kortlandt 2003).

In fact one of the oldest existing Slavic documents, the so-called Kiev Fragments, resemble more than anything else a Central Slovak-Eastern Serbo-Croatian transitional dialect. This area retains traces of the earliest splits in Proto-Slavic before its speakers left their homeland in Transcarpathia. The first splits were between South and West Slavic. This is why some lects in one or the other of these branches show remarkable connections with lects in the other branch (Kortlandt 2003).

Once again this harkens back to Central Slovak's links between South and West Slavic mentioned above.

Orava Slovak is a Central Slovak dialect that shows heavy Polish influence, and it is theorized to have Polish roots (Jesenský 2014).

Liptov Slovak is the Central Slovak dialect on which Standard Slovak is based.

Dolnozemske, Lowland, or Extraslovakian dialects are Central and Western Slovak dialects spoken outside Slovakia in other nations. There are speakers in the Pannonian Plain in Serbian Vojvodina (not the same as Pannonian Rusyn speakers), and in Southeastern Hungary, Western Romania, and the Croatian part of Sylvania.

Pannonian Slovak sounds like “old-time” Slovak that resembles Central Slovak from 100 years ago. Slovak speakers say the dialect sounds funny but is completely intelligible. The Slovak speakers are well-liked in Serbia and co-exist well with Serbs, but they keep their own identity and are not assimilated.

They go to their own mother tongue schools (Zušťák 2015) and have the possibility to do university education in their language. They have

a [rich life](#): there are 55,000 speakers, some towns are completely Slovak-speaking, there is a lot of media in their language and the culture is booming. It is [being influenced](#) by Serbo-Croatian. Most speakers are [bilingual](#) in Serbo-Croatian. The region is [multicultural](#), with Slovak towns mixed with Serbian and Hungarian towns.

Pannonian has [very limited intelligibility](#) of Czech. They [say](#) that they can sometimes understand slow and careful speech, but they have a hard time understanding normal speech, and sometimes they cannot understand it at all. Others say they can [barely understand](#) it at all. This sort of a description tends to correlate with an MI of 30%. Hence, Pannonian may understand 40% of Czech. However, they can read Czech well, even though some even [experience problems](#) in reading Czech.

Pannonian Slovaks speak an archaic Slovak that did not go through the changes that Slovak went through during the Czechoslovak period. Speakers say that Modern Slovak has a "Czech" flavor to it, no doubt due to this period. The language lacks a lot of the Slovak words for modern technological and cultural things. However, speakers say that some older Slovaks still speak exactly the same language they do, especially those from isolated or rural areas.

Romanian Slovak sounds funny to Slovaks but is fully intelligible to them (Zušťák 2015). It is being influenced by Romanian.

Hungarian Slovak is undergoing changes due to Hungarian. They are mostly monolingual in Slovak. Although they have passive knowledge of Hungarian, most don't speak it.

In addition to being influenced by the major languages of the lands in which they live, Extraslovakian dialects are experiencing changes due to long-term separation from Slovak Proper.

Slovak intelligibility of other Slavic lects:

Slovak intelligibility figures benefited from many Slovak informants, including an excellent personal informant.

Slovak has full 99% intelligibility of Pannonian Slovak.

Slovak has full intelligibility of Gemer Central Slovak at 95%.

Slovak has full 94% intelligibility of spoken Czech* (Golubovic and Gooskens 2015a), which is very high. Written is even better at a full 98%* (Golubovic and Gooskens 2015a). However, Pannonian

Extraslovakian Slovak, a Slovak dialect spoken in Slovakia for a century or more, has been separated from Czechoslovak for so long that it has only 50% intelligibility of Czech.

This is fascinating because it shows that true inherent Slovak-Czech intelligibility may be as low as 50% as has been controversially theorized. The full intelligibility figures revealed by scientific testing then would be all a result of the very heavy Czech bilingual learning that almost all Slovaks in Slovakia experience from the time when they are very young and they are exposed to Czech cartoons around age 6.

However, the truth is that this very high figure is mostly so high because of massive bilingual learning on the part of Slovaks. It would be difficult to find a Slovak with little exposure to Czech in order to test true Slovak inherent intelligibility of Czech.

Slovak has 81% intelligibility of Eastern Slovak, but that is more of a guess than anything else.

Slovak has 75% intelligibility of written Saris Eastern Slovak (Zušťák 2015).

Apparently other Slovaks often find Eastern Slovak impossible to understand.

Complicating matters further is the fact that the villages near the Ukrainian border where Eastern Slovak is spoken have quite a few actual Ukrainian and Rusyn residents and speakers (Zušťák 2015). The Rusyn speakers speak Lemko and Slovakian Rusyn, but Slovakian Rusyn is Eastern Slovak, not Rusyn.

Slovak has very high written intelligibility of Silesian at 80% (Zušťák 2015). We do know that Slovaks understand a high level of written Silesian. This makes sense as Silesian is between Polish and Czech, and Slovak has high intelligibility of Czech.

Slovak written intelligibility of Sulkovian Silesian is 70% (Zušťák 2015). Sulkovian Silesian is Silesian-Polish transitional, so it's a bit further away from Czechoslovak towards Polish than Standard Silesian is.

Slovaks understand Pannonian Rusyn quite well at 55%, but the sample size was very small. This is an Eastern Slovak dialect, and Slovak understanding of it is lower than for most Eastern Slovak

dialects. Pannonian Rusyn has been influenced heavily by Serbo-Croatian, so this explains the lower intelligibility.

Reports of decent Slovak-Polish spoken intelligibility are correct, with a figure of 50%* (Golubovic and Gooskens 2015a). Written is also very good at 59%* (Golubovic and Gooskens 2015a). However, the Southern Slovak dialect spoken down near the Hungarian border has very low understanding of Polish, only 10%. The Hungarian influence here is strong and may contribute to the very poor understanding figure.

Slovak has partial intelligibility of Rusyn at 45%. Written is better at 59%.

Slovak has much higher intelligibility of Western Ukrainian than for Standard Ukrainian at 40%.

Slovak has surprisingly high intelligibility of Čakavian at 40%. A connection between Slovak and South Slavic has long been theorized, but linguists sometimes scoff at it. Nevertheless, the high MI between Slovak and the older Croatian languages implies that there may be something to the Northern South Slavic-Slovak proposal after all.

Slovak has decent intelligibility of spoken Serbo-Croatian at 28%* (Golubovic and Gooskens 2015a). Written is a bit better at 37%* (Golubovic and Gooskens 2015a). Because Slovak and Serbo-Croatian intelligibility is decent enough, many Slovaks like to vacation in the former Yugoslavia, particularly in Croatia and Montenegro because they have coastal access (Zušťák 2015).

The closeness of Slovak and Serbo-Croatian at first glance seems mysterious, as the former is West Slavic, and the latter is South Slavic.

However, the connection is not so odd, and if we theorize that there was formerly a dialect continuum connecting West and South Slavic this makes more sense. The homeland of at least some of the Northern Serbo-Croatian speakers (Kaikavian speakers) seems to have been to the north in Central Slovakia. Slovene may have been connected to Czech, and Serbo-Croatian may have been connected to Slovak, in particular to Central Slovak (Nuorluoto 2010).

This theory states that there was a group called the Pannonian Slavs living in the land that is now Hungary. This group was dispersed to the north into Central Slovakia and to the south towards Croatia as

the Magyar Hungarians moved into the area. To the west, the connection between Slovene and Czech was disrupted when Austrian Bavarians moved into the area (Nuorluoto 2010).

Indeed MI is good between Slovak and Kaikavian and Čakavian, the remaining languages deriving from the Old Croatian that the first Croatian migrants from Central Slovakia spoke, adding an extra argument to this theory sometimes derided by linguists.

Slovak has 28% understanding of Old Church Slavonic.

Slovak has 27% intelligibility of Belarussian, higher than Ukrainian or Russian.

Slovaks understand Croatian at 25%. Croatian is closer to the old Croatian-Slovak connection, so it is better understood than Serbian.

Slovak intelligibility of Serbian is 20%.

Slovak has 20% understanding of Ukrainian. Written understanding is better at 35%. Note that this 20% figure is markedly lower than the 40% comprehensibility of heavily-Polonized Western Ukrainian.

Slovak intelligibility of Russian is very low at 16%. Written is better at 35%. Ukrainian and Belarussian comprehension is a bit better for Slovaks because each is rather Polonized and Slovaks understand Polish well.

Slovaks understand Slovene poorly at 14%* (Golubovic and Gooskens 2015a). Slovene writing is understood quite a bit better at 29%*. The poor understanding of Slovene does not bode well for a Slovak-North Southern Slavic connection. Instead the connection seems to be specifically with Croatian itself, specifically with Old Croatian, and not Northern South Slavic as a whole.

Slovaks have very low understanding of Bulgarian when spoken at 11%* (Golubovic and Gooskens 2015a). When written, understanding is quite a bit better at 25%* (Golubovic and Gooskens 2015a). In general, Slovaks have very poor understanding of Southern South Slavic, which is about as far away from Slovak as you can get.

Slovaks have even poorer intelligibility of Macedonian, another Southern South Slavic language, at 5%. Even worse, Slovaks understand nothing at all of written Macedonian.

Eastern Slovak

Eastern Slovak consists of the *Spiš*, Northeastern/*Šariš*, *Zemplín*, *Poprad*, *Preshov*, *Michalovce*, *Košice*, *Trebishov*, *Humenne*, and *Abov* dialects spoken in the eastern part of Slovakia. Eastern Slovak started to split away from the rest of Slovak very early. The first evidence of it appears as a threatening missive in 1493. Eastern Slovak has no Czech influence, but there is significant influence from Polish and German.

The character of Eastern Slovak has been somewhat controversial. A few linguists have suggested that Eastern Slovak is either Polish or Ukrainian that was subsequently Slovakized. However, the majority view at the moment is that Eastern Slovak is a Slovak language that came under heavy Polish and Ukrainian influence. In Eastern Slovakia and the neighboring Carpathians, Polish, Ukrainian, and Slovak have all been influencing each other for a long time, and in many cases, all three have come under Hungarian influence (Bidwell 1966).

It has [difficult intelligibility](#) with Western Slovak (M. Jembrih 2015). But the dialects that are difficult to understand for West Slovaks are typically in the rural areas, as big cities such as Kosice in Eastern Slovakia speak a Standard Slovak only slightly different from that spoken in big cities in Western Slovakia. Rural Eastern Slovak also has borrowings from German and East Slavic. Rural West Slovak dialects also have strange words that urban West Slovaks are not familiar with (Zušťák 2015).

Eastern Slovak is quite different from Standard Slovak. While in Slovak the stress is always on the first syllable, in Eastern Slovak it is always somewhere else, typically on the penultimate syllable. So two cognate or even identical terms will have [stress on different syllables](#) in Slovak and Eastern Slovak.

Although it will be quite controversial, a good case can be made that there is a [separate language](#) called Eastern Slovak. It is now mostly spoken in the rural areas, as the big cities have switched over to Standard Slovak. Eastern Slovaks have long rejected Standard Slovak books, saying that they were not written in the language that they speak.

One argument against Eastern Slovak as a separate language is that

while it is uncontroversial that Western Slovaks [cannot understand](#) Eastern Slovak, on the other hand, speakers of Central Slovak dialects (not Standard Slovak) are able to understand it (M. Jembrih 2015). So we are dealing once again with a dialect chain like so many others in the Slavic region.

In fact, some linguists describe all of Slavic as a huge dialect chain. That's not quite correct, but it's not completely wrong either.

In this work, strong dialect chains such as exist in the east and west of the Ukrainian and Lach languages and the north and south of the Kashubian language were initially the focus of much confusion, as I could not figure out what to do with them. Calling them all one language did not make sense, as the far ends of the chain could not understand each other, and unintelligible dialects cannot exist within a single language by definition.

Nevertheless, if you split the far ends into separate languages, you still end up with a central variety that is intelligible to everyone on the continuum. So while the ends are split well, the center of the continuum is now confusing, as we now have a language that is intelligible with two other languages, which should not be possible.

There's no effective solution here that solves all of the problems, so I decided to split the chains into three languages: 1) West, Central or Proper, and East, or 2) North, Central or Proper, and South, with the central language serving as the standard form of the macrolanguage.

If it worked well with Ukrainian, Lach, and Kashubian, this should also be an effective solution for East Slovak, especially as West and Central Slovak are not even part of a single language but are instead of dialects of a language called Czechoslovak in which the entire dialect chain is mutually intelligible. Obviously throwing East Slovak in as the only mutually unintelligible lect in a mutually intelligible dialect chain has problems.

Further, in an "intuitive" sense, East Slovak has the "look" of what we might call a language other than a dialect and this intuitive sense was heavily used for classification in this paper as it seemed to give better results than pure logic, which has its limits even in science, especially in the social sciences.

Šariš or Northeastern Eastern Slovak is spoken in Northeastern Slovakia near the Polish border and is closer to Polish and Ukrainian

than to Czech. It has [Polish roots](#) (Jesenský 2014). Slovaks say that Šariš is generally not understood in Slovakia outside of the eastern region. People who hear Saris often think they are listening to either Ukrainian or Rusyn. Šariš is close to both languages.

Zemplín Eastern Slovak has a long literary tradition. Zemplín has been written, generally liturgically, since the late 1700's. Speakers are Greek Catholics, as are most Eastern Slovaks.

Slovakian Rusyn, spoken in the Preshov Region of Slovakia, appears to be simply be a dialect of Eastern Slovak.

It resembles the Northeastern Eastern Slovak dialects such as Spiš, Šariš, Preshov, and Zemplín most of all (Vanko 1997). Slovakian Rusyn has influences mostly from Slovak but partly from Czech also (Kushko 2007). In recent years, it has borrowed heavily from Standard Slovak, especially in modern terminology such as legal, administration, and management (Vanko 1997).

Slovakian Rusyn is divided into two main dialects, *East Zemplín* and *West Zemplín*. Zemplín is also the name of an Eastern Slovak dialect. It is spoken in villages such as Benyadikovtsi, Tiglia, Krayna Polyana, Pristina, Shmygovets, Mikhailov, Novoselitsa, Ulych Tsintyr Pryslop, Kruzhliev Zvintir, Dubrava, and Gribov.

Slovakian Rusyn was finally codified into an official literary language in 1995.

The literary language is based on a mixture of the East Zemplín and West Zemplín dialects. From 1994-2002, a fair number of literary works were produced in Slovakian Rusyn. In addition, there is a Slovakian Rusyn newspaper (Vanko 1997).

Both the literary authors and the people who wrote for the paper frequently had a hard time finding the Slovakian Rusyn words for various things. In a number of cases, they had to substitute Ukrainian and Slovak words for concepts and items that Slovakian Rusyn had no word for. These Ukrainian and Slovak words were often changed in accordance with Slovakian Rusyn grammatical rules (Vanko 1997).

Part of the problem with the Slovakian Rusyn standard is that while the rules for the phonetics, morphology, and syntax of the standard were well-laid out, no process was created for the coining of new words via word formation principles and patterns of word derivation. Hence there seems to be no obvious way to coin new words in the

standard (Vanko 1997).

Slovakian Rusyn represents a classification dilemma, as it is intelligible with both Eastern Slovak and the rest of Rusyn. So is Slovakian Rusyn part of Rusyn, or is it part of Eastern Slovak? Slovakian Rusyn seems very close to Pannonian Rusyn, which is clearly part of Eastern Slovak and not a part of Macro-Ukrainian Rusyn Proper (Lund 1997). Since Slovakian Rusyn seems to be the same language as Pannonian Rusyn, if Pannonian Rusyn is Eastern Slovak, then Slovakian Rusyn is also. It is best to put Slovakian Rusyn, like Pannonian Rusyn, outside of Rusyn Proper and into Eastern Slovak instead.

Slovakian Rusyn is reportedly intelligible with Eastern Slovak, as speakers of the two dialects [readily communicate](#) in their own dialects with each other in the towns in Northeastern Slovakia where they both reside and in mixed families where some speak Rusyn and others speak Eastern Slovak (Vanko 1997).

In Slovakia, Slovakian Rusyn-Eastern Slovak communication has been called semi-communication (Vanko 1997).

Semi-communication is a concept first developed by Einar Haugen in 1966 to describe communication between speakers of different Scandinavian languages (Haugen 1966). It does not necessarily mean that only half the conversation is understood as one might expect from the term. Instead it simply refers to a process where speakers of two different dialects or languages communicate with each other, each in their own language, and the process goes fairly well (Kaplan and Baldauf 2005).

However, when they are outside of Northeastern Slovakia, Slovakian Rusyn speakers use Standard Slovak to speak to Slovaks, as all of them are Standard Slovak speakers (Vanko 1997). While Eastern Slovak speakers may understand Slovakian Rusyn well, Standard Slovak speakers have poor intelligibility of Slovakian Rusyn.

Slovakian Rusyn is reportedly intelligible with Rusyn dialects spoken in other countries (Plishkova 2008).

Plays were performed in Ukrainian in the Slovakian Rusyn region for 30 years. During that entire period, the Slovakian Rusyns complained that they could not understand the plays (Vanko 1997). The fact that they never learned to understand the plays over decades implies that

Eastern Slovak and Ukrainian may not be as close as so many think. Furthermore, this data seems to sound the death knell for the standard view that Slovakian Rusyn is a part of Rusyn, and all of Rusyn is a dialect of Ukrainian. This is one more piece of evidence indicating that Rusyn is a separate Slavic language, part of Macro-Ukrainian certainly, but a Ukrainian dialect, certainly not.

Vojvodinian, Pannonian, or Barachka-Srem Rusyn is spoken by a group of Rusyns who migrated to Northwestern Serbia (the Bachka region in Vojvodina province) and Eastern Croatia from Eastern Slovakia and Western Ukraine 250 years ago. There are only 20,000 speakers remaining. It is mostly spoken around Novi Sad and Kula.

It has since undergone heavy South Slavic influence from Serbo-Croatian (Martin 2013). It also has influences from Hungarian (Kushko 2007). Prior to WW1, the region was under the control of the Austro-Hungarian Empire, and the prestige language in the area was Hungarian. Many Hungarians continue to live in Vojvodina to this day. The Pannonian Rusyns borrowed many Hungarian terms directly from the Hungarian speakers they lived with. There are also German and OCS loans (Bidwell 1966).

There is a question about whether the Hungarian, German, and OCS loans entered Pannonian Rusyn in a roundabout way via Serbo-Croatian or whether they came in straight from Hungarian, German, and OCS. The available data suggests that they came in straight from Hungarian, German, and OCS (Bidwell 1966).

However, a couple of the Hungarian loans seem to have entered the language in the Pannonian Rusyn homeland in Eastern Slovakia, an area which has long been under strong Hungarian influence (Bidwell 1966).

Pannonian Rusyn is an official language in the Vojvodina region (Martin 2013). It was first codified in 1923 (Vanko 1997).

An official standard was created for Pannonian Rusyn for use in education and publishing in 1962 and has been in use ever since (Maksimovich 2008). They have their own schools where they learn in their mother tongue.

Ukrainian-oriented linguists have long said that Pannonian Rusyn is a Ukrainian dialect like Lemko and Carpathian Rusyn (Lund 1997). This is not at all the case.

Recent studies sealed the matter, proving that Pannonian Rusyn is simply an Eastern Slovak dialect transplanted to Yugoslavia. It has many characteristics of Eastern Slovak that Ukrainian lacks, and every feature that it is similar to Ukrainian it shares with Eastern Slovak. So it is all Eastern Slovak and not Ukrainian at all (Lund 1997).

The matter is complicated because Pannonian Rusyns regard themselves as Ukrainians and say they are not Slovaks even though they came from Eastern Slovakia and speak an Eastern Slovak dialect. This shows that ethnic identity is not always the same as linguistic identity (Lund 1997).

The homeland of the Pannonian Rusyns was probably the Slovakian Rusyn area around Preshov. These are Rusyns who also call themselves Ukrainians but live in Eastern Slovakia and speak Eastern Slovak. So the Pannonian Rusyns may be Slovakian Rusyns transplanted to Yugoslavia. As Slovakian Rusyn is for all intents and purposes a part of Eastern Slovak, it makes sense that Pannonian Rusyn is also a part of Eastern Slovak.

One problem is that there is no information on Eastern Slovak-Pannonian Rusyn mutual intelligibility. The working hypothesis is that it is full, as the Pannonians have only been in the Balkans for 250 years, which in many cases is not long enough to form a separate language.

However, in some cases, separate languages can develop after a separation of only 100-200 years, as in the case of Chinese Kazakh-Standard Kazakh, where Chinese Kazakhs can no longer understand Kazakh broadcasts because over the last 100 years, many Russian loans have gone in to Kazakh, and in the case of Turkish Crimean Tatar-Crimean Tatar, where 200 years separation have caused many Russian loans to go into Crimean Tatar and many Turkish loans to go into Turkish Crimean Tatar, resulting in them no longer being mutually intelligible.

Furthermore, many Serbo-Croatian loans have gone into Pannonian Rusyn over that 200 year period, and Slovak speakers now have much poorer intelligibility of Pannonian Rusyn than of Eastern Slovak. Nevertheless without information about the MI between these two Eastern Slovak dialects, we are unable to separate them into separate languages.

Intelligibility between Pannonian Rusyn and the rest of Rusyn is unknown.

Pannonian Rusyn intelligibility of other Slavic lects:

There was a dearth of Pannonian Rusyn informants which made MI study difficult.

Pannonian Rusyn has 55% intelligibility of Serbo-Croatian.

Eastern Slovak intelligibility of other Slavic lects:

There are serious problems with the intelligibility figures for Eastern Slovak. A lack of Eastern Slovak informants seriously hampered the study of this language, so figures should be taken with a degree of skepticism.

Eastern Slovak has 50% intelligibility of Rusyn. Reports show that Eastern Slovak has some intelligibility of Polish and Rusyn. However, Slovakian Rusyn (Eastern Slovak) speakers say that they can only understand 50% of Rusyn. The divide between Eastern Slovak and Rusyn, as between all branches of West and East Slavic, is sharp.

Eastern Slovak understanding of Ukrainian may be 45%. Eastern Slovak intelligibility of Ukrainian is not as good as with Rusyn. Slovakian Rusyn speakers in Slovakia say they cannot understand plays performed in Standard Ukrainian (Vanko 1997). As Slovakian Rusyn is an Eastern Slovak dialect, this evidence shows that Eastern Slovak is not intelligible with Standard Ukrainian.

Eastern Slovak intelligibility of Polish may be 40%. Eastern Slovak comprehension of Polish is less than that of Ukrainian.

Eastern Slovak has 40% intelligibility of Kaikavian. Once again we see the remarkable closeness between Slovak and the remains of Old Croatian.

Eastern Slovak may have 30% intelligibility of Russian. Eastern Slovak intelligibility of Russian is less than of Polish.

Czech

Czech is part of the *Czechoslovak* macrolanguage which consists of four languages: Czechoslovak with two dialects, Czech and Slovak,

Eastern Lach, Western Lach and Eastern Slovak.

Czech actually a macrolanguage consisting of two or possibly four languages: Czechoslovak, Western Lach, Eastern Lach, and Lach. At least one of these, Lach, is a certainly a separate language.

Although it will be quite radical, in this work I will propose that the Czech and Slovak languages simply do not exist. Instead, they are dialects of a single tongue, Czechoslovak. As this study splits based on intelligibility alone, and Czech and Slovak have 91% and 94% intelligibility of each other proven by formal study (Golubovic and Gooskens 2015a), there is no linguistic basis to split these two lects into separate languages. Certainly there are political and sociological reasons to split them, but this study does not deal with those and only splits or lumps based on MI.

Czech is something of an artificial creation, which is true of most Slavic literary languages. It was created by a committee around 1900 (Kaplan and Baldauf 2005) based on a mixture of all of the existing Czech dialects. The standard, like Italian based on the language of Dante, was based in part on the standard literary Czech from the 1600's (Rusinko 2003). Most Czechs speak a regional dialect, not Standard Czech, in their day to day lives. Standard Czech from the 1600's and Standard Slovak from the 1800's are fairly close to each other, closer than many Czech or Slovak dialects are to each other.

Under Austro-Hungarian rule, the Czech language was severely repressed in favor of German (Cerna and Malchalek 2007), and it is almost fortunate that it survived.

Czech has very good intelligibility with Slovak, but nevertheless the extent of this intelligibility is much exaggerated. It is true that Czech has good intelligibility of Western Slovak dialects, but that is not necessarily true of Central, Eastern, and Extraslovakian Slovak lects.

The MI of Czech and Slovak, not to mention completely separate languages like Czech and Eastern Slovak, is wildly controversial, with some people, especially Slovaks, insisting that everyone understands everyone perfectly and others stating that MI is a considerably more problematic than that, mostly in terms of Czech understanding Slovak and not the other way around.

Since the breakup, young Czechs and Slovaks understand each other worse since they have less contact with each other. Czechs aged 20-

35 grew up with little exposure to Slovak. Younger Czechs got more exposure as Slovak news broadcasts and game shows were added to Czech TV in the past ten years. Even Czechs who understand Slovak well say they would not even try to speak it.

In Czechoslovakia, even many older Czechs are losing their knowledge of Slovak to the point where Czech understanding of Slovak in Czechoslovakia is becoming difficult, and Slovaks are becoming "othered" in the country as in "not from here."

Recent reports from Czechs say that they understood Slovak well 10 years ago, but in the past 10 years, the languages have been growing so far apart that they no longer understand it so well.

Formal studies found that Czechs and Slovaks have 91-94% intelligibility of each other's language (Golubovic and Gooskens 2015a), but some Czechs report that young people often have ~80% intelligibility of Slovak because there is so much less communication between them nowadays as opposed to when they were one country as Czechoslovakia.

In 2005, there was a bit of a scandal after a Slovak show aired on Czech TV was given subtitles.

In the former Czechoslovakia, everything was 50-50 bilingual - media, literature, etc. Since then, Slovak has been disappearing from Czechia, so some of the younger people don't understand Slovak well.

Broadcasters under Communism used a very correct, soft form of Slovak that is not often heard today (Zušťák 2015). However, since the split, the languages have been drifting further apart, and changes between the two have been accelerating.

The two languages are often said to be as far apart as British English and US English, but this is not correct, as the differences are quite a bit greater than that. The grammatical and lexical differences between Czech and Slovak are far greater than between the two Englishes. A better comparison would be between Swedish and Danish or Scots and English – close, but certainly not two derivatives of one language. A better comparison to the two Englishes would be the difference between Moravian Czech and Bohemian Czech.

Czech intelligibility of Slovak is 91% (Golubovic and Gooskens 2015a). However, this high figure is surely due to bilingual learning.

Some dialects do have excellent inherent intelligibility.

For instance, *Moravian Czech* (Eastern Czech) and *Bratislavan Slovak* (especially *Moravian Slovak*) speakers communicate very well. The dialects are quite different, but they have excellent intelligibility (Zušťák 2015).

It is hard to find the "virgin ears" necessary for intelligibility testing of Czech and Slovak, as Czechs and Slovaks have so much exposure to each others' languages (Zušťák 2015).

Inherent intelligibility of Czech with Slovak may be as low as 50% according to some reports, as some say that young Czechs with little exposure to Slovak have intelligibility this low. Some Czech parents say their children tell them that they do not understand Slovak. This may be true of younger children, but by teenage years, young Czechs seem to understand Slovak well.

The 50% figure for Czech-Slovak inherent intelligibility is controversial, and some Slovaks strongly disagree (Zušťák 2015).

Intelligibility problems are mostly on the Czech end because they don't bother to learn Slovak, while all Slovaks learn Czech (Kaplan and Baldauf 2005). There is as much Czech literature and media as Slovak literature and media in Slovakia – many books in Slovakia are only available in Czech (Kaplan and Baldauf 2005) - and many Slovaks study at Czech universities. When there, they have to pass a language test. Czechs hardly ever study at Slovak universities. So there is quite a bit more motivation of Slovaks to learn Czech than the other way around.

In recent years, there has also been quite a bit of bilingual learning among Czech and Slovak speakers. Young Czechs and Slovaks talk to each other a lot via the Internet (Zušťák 2015). There are also [some TV shows](#) that show Czech and Slovak contestants untranslated (like in Sweden where Norwegian comics perform untranslated), and most people seem to understand these shows.

Many foreign movies get translated into both Czech and Slovak. People tend to prefer one language version of the movie over the other, depending on the movie. The Czech movies often try to translate all the foreign names of characters, often with what Slovaks consider to be hilariously bad results, whereas movies translated into Slovak often maintain the original names of characters (Zušťák 2015).

There has been some discussion of Czech-Slovak MI in the Linguistics literature (Nábělková 2007, Budovičová 1987).

The discussion about Czech-Slovak MI doesn't seem to be going anywhere. Sober-minded discussions by mature people can be found on the Net, but so can passionate hatefests. The formal studies of the standard dialects have already been done, so that matter is seems settled.

But the larger questions about MI on the more micro-levels of individuals and through time is still up in the air, and the discussion about MI between dialects and even microlanguages like East Slovak continues to rage unabated. Studies could be done but there are few funding prospects and not much interest in interdialectal MI studies. Even studies would probably not settle much. This debate is likely to rage on into the foreseeable future with little reconciliation in sight. Czechs and Slovaks are very civilized as humans go – the split between states was one of the most amicable in modern history – but these people are emotional mammals like all the rest of us humans.

Czechs see Slovaks as country bumpkins - backwards and folksy but optimistic, outgoing, and friendly. Czechs are more urbane. The written languages differ much more than the spoken ones.

The languages really split about 1,000 years ago, which should make them as further apart than Spanish and Portuguese and as far apart as English and Old English.

However, written Slovak was based on written Czech (Kaplan and Baldauf 2005). There was also a lot of interlingual communication.

It is estimated that 80% of Czech and Slovak vocabularies are identical or similar. One study found that out of the 500 most commonly used lexemes or words, 230 were identical between the two languages, 154 were either similar or related, and 116 were completely different (Kaplan and Baldauf 2005). However, 22% of the most commonly used words being completely different is a pretty significant distance between lects. However, the grammars of the two languages are quite similar: nouns and verbs are inflected in both languages.

Some letters are different: *ř*, *ě*, and *ů* occur only in Czech; and *ä*, *dz*, *dž*, *í*, *ľ*, and *ŕ* and the diphthongs *ia*, *ie*, *iu*, and *ô* occur only in Slovak. The Czech *ř* is hard Czech second language learners to pronounce,

and Slovaks have difficulty with it too. It only occurs in one other language on Earth, and its occurrence in even that language is controversial.

Two of the Slovak consonants, *í* and *ř*, are actually long consonants. These consonants (and their short forms also) can actually serve as vowels in syllables, a source of difficulty to anyone who learns Slovak as a second language. The result is words that appear to lack vowels altogether such as *stíp*, *víba*, *smršť*, or *zmrzlina*.

Slovak has six noun cases, while Czech has seven. The extra case for Czech is the vocative case, which is almost extinct in Slovak. For instance, Czech nominative *Petr* becomes vocative *Petře* when Peter himself is being spoken to.

Literary Czech is a written language that is never spoken.

But spoken Czech definitely differs from written Czech. There are two varieties of spoken Czech. One is Standard Czech. This is mostly used on TV and for official functions. Even ordinary citizens when being interviewed on TV will try to speak Standard Czech. On the other hand, the normal everyday speech on the streets, at work, and in the homes is a Colloquial Czech that is a much less formal version of Standard Czech. It is spoken across the land that has significant differences in accent, a few different words and small differences in pronunciation. Young Czechs tend to speak the same way across the country.

The dialects have been heading out in Czechia for a long time and are now mostly spoken by the elderly. They started going extinct 80-90 years ago. Dialect leveling has occurred for a variety of reasons, including heavy urbanization, a recent centralizing tendency in the nation, the lowland nature of most terrain, and the fact that the urban-rural linkage was never badly severed here. Almost no Czech young person speaks dialect, and the very idea that one would seems ludicrous. The dialects are seen as either humorous, boorish, or a sign of lack of education, and these views have no doubt helped to drive them towards extinction.

The leveling trend may go back even further than a century ago. The Czech literary language nearly went extinct due to Germanization in the 1600's-1700's. The panic over losing their language was one of the reasons for the Czech independence movement.

But literary Czech had been diminished so badly that the makers of the new literary language being created around 1850 thought it was too far gone to be rescued, so they reached back for a 1500's Czech to base the new Literary Czech on. Obviously, this archaic Czech was so foreign to Czechs that it had to be nearly taught anew as a nearly foreign language. The recent imposition of a popular literary standard may have jump-started that Czech dialect death that began in earnest in the 1920's.

The leveling trend seems to have effected other languages too. 100 years ago, varieties of German, Slovak, Hungarian, Rusyn, and Romani were widely spoken across Czechia. The Germans mostly left the Bohemian region after the war for obvious reasons and there are probably few if any Rusyns left.

Bohemian Czech (Western Czech) is spoken in the western region of Bohemia. While most Czech dialects have changed dramatically in the last 100 years, mostly by heading towards extinction, it is interesting that Bohemian has changed little in the last century and the Bohemian of 100 years ago looks a lot like colloquial Czech. Recently the colloquial Czech that has spread across the land seems to be based on the Bohemian dialect.

Moravian Czech (Eastern Czech) is quite different from the rest of Czech, and there is a question over whether it is a separate language from Standard Czech. As things stand, it is the most divergent remaining dialect in Czechia. There is a possibility that some Moravian Czech speech may not be well understood in the rest of Czechia, as some Moravian speakers try to speak Standard Czech for fear that their Moravian speech will not be understood. Moravian and Bohemian dialects used to be almost identical, but they began diverging in the 1500's, especially in vowel pronunciation.

However, the project to have Moravian declared a separate language has so far been run by crackpots and nonprofessionals and has been dismissed by linguists as not serious. There is no good evidence that Moravian is a separate language from Czech.

Eastern Moravian is close to Western Slovak.

Northern Moravian is close to Polish.

Central or Hanák Moravian encompasses most of the Moravian area.

All Moravian dialects are now in retreat and are being replaced with a

Bohemian-based dialect on which Moravian has left few traces.

Despite the fact that dialects are dying out in Czechia, few Czechs actually speak Standard Czech in their day to day lives, so, like Standard Croatian, we are dealing once again with a standard language that almost nobody uses in their quotidian lives. One of the Northeastern Bohemian is said to be very close to Standard Czech, but in most of the country, people continue to use watered down versions of the old Czech dialects.

Czech intelligibility of other Slavic lects:

The study of Czech MI benefited from many Czech informants; however, there were no personal informants.

Czechs have 94% intelligibility of oral Slovak* (Golubovic and Gooskens 2015a) and 97% intelligibility of written Slovak* (Golubovic and Gooskens 2015a). Czechs generally claim that they can understand written Slovak completely, and this claim was proven correct via scientific testing.

Czech intelligibility of Pemci Czech spoken by Czechs who have lived in Croatia for centuries is good but not full. A decent figure may be 80%.

Czech has 80% intelligibility of Gemer Central Slovak. As you can see, individual Slovak dialects get harder to Czechs to understand as they move towards the east.

Czech intelligibility of Eastern Slovak is estimated at 68%. Czech has limited intelligibility of Eastern Slovak. This is confirmed by Slovak speakers, even those who take issue with Eastern Slovak as a separate language. Even they say that Western Slovaks cannot understand Eastern Slovak, so obviously Czechs can't either.

Czechs understand Pannonian Rusyn fairly well at 70%. This is because this language is actually part of Eastern Slovak and not Rusyn. The 70% figure for Pannonian Rusyn is probably generalizable to Eastern Slovak also.

Czechs have 52% intelligibility of Upper Sorbian. Czechs also report 100% understanding of written Upper Sorbian. This high figure seems strange. However, oral intelligibility figures were all over the place, with some saying that they understood it completely and others saying they had little understanding of the language. Czech

understanding of Upper Sorbian is much higher than the other way around, which is odd.

Czech intelligibility of Lach is not known but is probably lower than Polish understanding of Silesian as there more distance between Czech and Lach than between Silesian and Polish. A good figure is not available, but a good estimate might be ~50%.

Czechs report only 50% intelligibility of Pannonian Extraslovakian Slovak, an archaic form of Slovak.

Czechs have 50% comprehension of Serbian. This figure is much higher than for Serbo-Croatian itself, which is odd as Serbian is simply a dialect of Serbo-Croatian, so this figure doesn't make much sense.

Czech intelligibility of Cieszyn Silesian is estimated at 37%. This figure is very rough and is derived from large intelligibility range. Czech understanding of this language can definitely be very poor and in some cases has been described as zero. A Slavic enthusiast on the Net gave a figure of 75%, but that's probably more of a guess than anything else. So the intelligibility range may be 0-75%. Taking that figure in half gives us 37%, about the same figure as for the MI between Serbo-Croatian and both Kaikavian and Čakavian.

Czechs have 36% intelligibility of spoken Polish* (Golubovic and Gooskens 2015a), about the same as the other way around. Czech written intelligibility of Polish is a bit better at 46%* (Golubovic and Gooskens 2015a). It is often said that Czech is intelligible with Polish, but this is just not true. Conversational attempts between Czechs and Poles typically have poor results, and the two groups do not understand each other well enough to have a normal conversation, but some basic level communication is possible.

Czechs and Poles who meet often use English to talk to each other rather than bother with the vagaries of Czech-Polish miscommunication. Czechs say that when Poles speak, it sounds like they are speaking an archaic version of Czech similar to the Middle Czech spoken in the Middle Ages.

Older Czechs understand Polish much better since they were exposed to a lot of Polish cartoons when young. Much of the claimed Czech-Polish intelligibility is simply bilingual learning. A Czech speaker who also speaks Russian will find Polish easier to understand. As in Romance, the more Slavic languages you know, the better your

understanding of other Slavic languages will be. However, Czechs can learn Polish fairly easily. If a Czech moves to Poland, he may have good Polish within a year of the move.

Czech understanding of Rusyn is 32%. This is much lower than the Slovak figure.

Czech has poor oral understanding of Slovene at 18%* (Golubovic and Gooskens 2015a). Written intelligibility is only a bit better at 27%*.

Czech has 27% understanding of Ukrainian. The Slovak figure is higher.

Czech oral intelligibility with Slovene is 18%* (Golubovic and Gooskens 2015a). Czech understanding of written Slovene is a bit better at 27%* (Golubovic and Gooskens 2015a).

Czechs understand Serbo-Croatian at 18%, the same as with Slovene. However, written comprehension is much higher at 45%, which seems a bit odd.

Czechs have only 15% understanding of Bosnian, but this figure is along the lines of that of Serbo-Croatian, so it seems about right.

Czechs have very low understanding of Russian at 13%. The other way around is low too, so Czech-Russian MI is simply poor. However, Czechs understand written Russian very well at 75%, which doesn't make much sense.

Higher oral figures are sometimes reported, but they can be confounded by a great deal of bilingual learning. For instance, all Czechs over age 35 went through mandatory Russian classes under Communism. They had to take up to 12 years of Russian. Some did well and enjoyed their Russian classes, while others suffered through them. All of these people can speak Russian at least somewhat, but they do not like to, and the Russian they speak is often broken. Young people typically cannot understand Russian at all.

Czechs have a poor understanding of Bulgarian speakers at 13%* (Golubovic and Gooskens 2015a). Understanding of Bulgarian writing is also low, 19%* (Golubovic and Gooskens 2015a). The far end of South Slavic is very far from the Czech language.

Czechs understand Macedonian very poorly at 13%, the same as for Bulgarian, which makes sense.

Czechs can hardly understand Kashubian at all at 9%. This is strange because Kashubian is another West Slavic language.

Czech understanding of Belarussian is very poor at 5%. Czech understanding of East Slavic in general is quite poor.

Pemci Czech

Pemci Czech or Croatian Czech is an interesting split because I have never heard of anybody splitting this Czech group off. The main point to keep in mind though is that these Czech people did not show up yesterday. Or even 50 years ago. Or even a century ago. They have been here a very long time, long enough for their Czech to change dramatically.

The Czechs in Croatia are called Pemcis by non-Czechs. The name is apparently not pejorative. The Pemci Czechs are a little known group of people. Indeed, in late middle age, I am just hearing about them for the first time. One problem is that the Pemcis are a small group of less than 10,000 people. The last census found 9,641 Pemcis in Croatia.

The Pemcis live in together in a compact group of their peers in part of Western Slavonia around the cities of Daruvar and Grubišno Polje, mostly in Bjelovar-Bilogora County and to a quite lesser extent in Požega-Slavonia County. They make up 5.25% of Bjelovar-Bilogora County and .83% of Požega-Slavonia County. They are a plurality in the city of Končanica and are also found in Veliki Zdenci, Mali Zdenci, Golubinjak, and other villages nearby. There are also very small numbers of Pemcis in all major cities in Croatia. Czech is now used as an official language in one city and in four villages.

Around 1700, the Ottomans surrendered the Slavonia region to the Hapsburg Austro-Hungarian Empire. There were many Muslims living in the area, and they all fled because after the Turkish wars, returning Christians sought vengeance against the Muslims for the wars the Turks waged on them. In one area, Christians burned 200 mosques to the ground.

The Muslim dispersal left vast swathes of empty land and the

Hapsburgs set to work settling the fallow lands with people from all over the empire.

The first Pemcis came to this part of Slavonia in 1750 and they continued to arrive all through the 1800's. Land was very cheap, selling in Slavonia at 10% of the price of land in Czechia.

Pemci Czechs also moved into Gorski-Kotar and all of the major cities where they were praised as excellent workers, but these people were all assimilated as Croats with two or three generations.

Other Pemci Czechs tried to retain their culture by organizing Česka Besede or Czech cultural organizations that tried to preserve Czech language and culture. Soon Česke Besede were organized all over the country. Libraries and sports associations sprung up. Soon after the turn of the century, the first Czech newspaper was being published in Zagreb.

The Proto-Yugoslavia state called the Kingdom of Serbs, Croats, and Slovenes founded after WW2 was very friendly to Pemcis. In the 1920's, Czech schools and kindergartens were built. Pemcis organized a Czech political party in the early years of this new state.

Most of both Czechs and Slovaks did not collaborate with the Axis Powers when Germany occupied the Kingdom. Many Pemcis fled back to Czechoslovakia after the war, however, as the Communists seized power in the new Yugoslavia, some of them returned to Croatia. In Communist Yugoslavia, Pemcis saw their rights dramatically expanded. Many more Czech schools were built.

In the Balkans War, Slovenia was a front line, and many Pemcis volunteered for the Croatian army.

Czechs are recognized as an official minority in Croatia and the Czechs and Slovaks combined are allowed to elect one representative to Parliament. In recent years, the Czechoslovakian representative has been from the Croatian People's Party.

There are now 24 Česka Besede all over the country. They organize dances, poetry readings, concerts, plays, festivals, and Czech language classes. There are now dozens of Czech kindergartens and primary schools in Croatia, and one high school in Daruvar has a Czech Department. Daruvar is a city with a large Pemci population.

The Pemcis also speak a peculiar variety of Czech. The Pemcis have

been in this region for 250 years and their Czech is archaic and sounds strange to modern Czechs. All or almost all Pemcis have spoken Serbo-Croatian for quite some time, as bilingualism has been the norm for this group. The long-term bilingualism resulted in Pemci Czech being. Czechs who listen to Pemci speech say that the large number of loans means they [cannot understand](#) the Pemcis at times. Although no figures have been gathered, that can mean intelligibility of 80%. This is also a typical MI figure for 250-350 years separation of two lects. Since Czechs lack full intelligibility of Pemci, it must be a separate language. Its archaic character and the many Serbo-Croatian loans also mean that the language is now quite a bit different structurally than Czech.

Lach

Lach, Lachian or Czech Silesian is a language made up of at least one or possibly two languages, **Eastern Lach** and **Western Lach**. The two geographical varieties of Lach in the west and east have [difficult intelligibility](#), hence a good case can be made that they are separate languages on linguistic grounds.

However, in recent years, Eastern Lach and Western Lach have been dying out, as a Lach koine has developed in Lach that can be understood by all Lachs (Hannan 1996b).

This seems to be a case of two languages gradually merging around a single koine.

Western Lach includes the area around Opava, while Eastern Lach encompasses the Ostrava region. In recent years, many Czechs from other parts of Czechia have moved into the Ostrava region (Hannan 1996b).

The dialect of the Zaolzie district has been characterized by some as East Lachian, though its phonology and morphology resemble Cieszyn Silesian (Hannan 1996b).

Lach is a Czech-Polish transitional language with a close relationship with Cieszyn Silesian spoken in northeastern Czech Republic to the east of the Cieszyn Silesian zone (Hannan 1996a).

In this region it is spoken in Northeastern Moravia, the region called Old Austrian Silesia, and the Hluchin district. The speaker area includes the Frydek and Teschen districts (Hannan 1996b).

Czech linguist Jan Balhar estimates that there are currently 300,000 Lach speakers (Hannan 1996b).

In 1934, Óndra Łysohorsky, a famous Czech poet who often wrote in Lach, estimated that there were 2 million Lach speakers. However, this figure includes all speakers of Cieszyn Silesian, Silesian, and possibly even Podhale (Hannan 1996b).

The term Lach dates back to the 1500's and was traditionally used as a pejorative term by highlanders for lowland agriculturalists. The term was first used by Wallachian shepherds who arrived in the Beskid Mountains to the east of the Lach region in the late 1400's and 1500's (Hannan 1996b).

These Wallachians later went on to form the core of the Lemko Rusyn speakers.

It is related to the West Slavic term *Lech*, used for Poles. The term *Lech* can be found in the name for the Polish languages, Lechitic (Hannan 1996b).

In the early 1900's, speakers of German, Polish, Czech, Slovak, Lach and Moravian Czech, thoroughly mixed together, were all common in the Lach region. During this period, Lach speakers claimed Czech, Silesian, Polish, Moravian and even German ethnicity (Hannan 1996b).

Even as of 1920, there was little Czech national consciousness in the region. The first Czech school in the region opened in 1918. In the 1920's, most of the schooling in the area was in German. German was the main language of government in the 1800's when this area was part of Austrian Silesia (Hannan 1996b).

Lach has been undergoing independent development outside of Czech since the 1200's-1500's (Hannan 1996b).

It is often considered to be a dialect of Czech, but this cannot be true. Nevertheless, even if Lach is seen as a Czech dialect, Lach is certainly the furthest removed from any of the standard Czech dialects and from Standard Czech (Hannan 1996b).

Lach is as close to Slovak as it is to Czech. Lach is particularly close to

Western Slovak. It is as close to Western Slovak as to any Czech dialect or to Standard Czech. It also shares a number of features with Eastern Slovak (Hannan 1996b).

Lach is very close to both Cieszyn Silesian and Podhale, and in fact it is sometimes confused with either of them (Hannan 1996b).

The differences between Cieszyn Silesian and Lach are uncertain.

In some cases, people seem to be referring to the same lect when they speak of Lach and Cieszyn Silesian (Hannan 1996b).

Despite its close relationship with Cieszyn Silesian and Podhale, Lach is better seen as a separate language, as it is not even fully intelligible within itself.

Another word for Lach is Czech Silesian, and some see Lach as simply Silesian as spoken in Czechia, albeit with massive Czech influences. However, Lach is not considered to be part of the Silesian language, so Silesian-Lach intelligibility is certainly not full. Lach is not fully intelligible with Czech; indeed, the differences between Lach and Czech appear to be greater than the differences between Silesian and Polish.

A look at comparison sentences between Czech, Polish and Lach shows that Lach is dramatically different from Czech, so different that it seems dubious that a Czech speaker could understand it (Hannan 1996b).

This is in spite of the fact that Lach has been heavily leveling into Moravian Czech for the last 100 years. In the last 100 years, Lach has changed so dramatically modern Lach speakers have a hard time reading the prose of writers who wrote in Lach in the 1920's-1940's (Hannan 1996b).

Łysohorsky was also a competent linguist, and he created an alphabet for the literary Lach language. The orthography consisted of one half Polish letters and one half Czech letters, in tandem with Łysohorsky's belief that Lach was midway between the two languages (Hannan 1996b).

Throughout his life, Łysohorsky insisted that Lach was an independent language and not a Czech dialect. He included Cieszyn Silesian and possibly Podhale in this independent Lach language (Hannan 1996b).

Czech linguists are unanimous that Lach is a part of Macro-Czech

(Hannan 1996b).

Polish linguists have three different views on Lach. Some say that Lach is indeed Lechitic, albeit with strong Czech influence. Others say that Lach is neither Polish nor Czech but instead transitional between the two languages. Yet others feel that Lach is Czech with significant Polish influence (Hannan 1996b).

The standard view among linguists seems to be that Lach is a part of Czech. The best analysis seems to be that Lach is part of Czech.

Certainly Lach has some features similar to Polish in the phonological system, especially the consonants (Hannan 1996b).

Formally, Lach is called Moravian-Silesian Czech. It is spoken in the parts of the Moravian-Silesian region that were not dominated by Russian speakers and are not part of the transitional Czech-Polish zone in the east where Cieszyn Silesian is spoken. In Czechia, Lach and Cieszyn Silesian are together called the Silesian dialects.

Lach is in similar shape as many Czech lects. It is spoken mostly by people middle-aged and older in villages in the rural areas.

Furthermore, all Czech lects have been weakening as the dialects merge and become part of broader interdialectal systems.

Even as early as 1979, there were no remaining monolingual Lach speakers, and all speakers of Lach now speak a form of Standard Czech also (Hannan 1996b). As all Lach speakers also speak Czech, Lach intelligibility of Czech is impossible to measure.

Sorbian

Sorbian is a macrolanguage consisting of two languages, Upper Sorbian and Lower Sorbian, which are spoken by small groups of people called Wends in the Lusatia region of Eastern Germany. The Sorbian languages have, like Slovene, retained the unusual dual number.

The Sorbs were originally from Ukraine and in the 500's moved from there to Germany. The Sorbian languages still contain many Ukrainian words.

The Sorbian languages are highly diverse, and one theory is that they are transitional between Northwestern Slavic Lechitic and Southwestern Slavic Czechoslovak (Sussex and Cubberley 2011, p. 55). This theory says Lower Sorbian is more Lechitic (Polish), and Upper Sorbian is more Czechoslovak. In this analysis, Lower Sorbian is Northwestern Slavic, and Upper Sorbian is Southwestern Slavic (Sussex and Cubberley 2011, p. 55).

Some evidence in support of this theory comes from an intelligibility study that showed that Upper Sorbian speakers understand Czech much better than they understand Polish, though they did not understand either language well (Sloboda and Brankaikec 2012). This implies that the Sorbian languages are more similar to the Silesian and Lach languages than to Polish itself.

People listening to Upper Sorbian say it sounds like a mixture of Czech and Polish spoken with a heavy German accent. This is the only Slavic language with a strong German accent. It's quite normal though, as Sorbs have been under the influence of the German language for 1,300 years.

The strong Czech influence makes sense because Lusatia was part of the Czech and Greater Moravian Kingdoms for centuries.

Some linguists regard the question of whether Lower and Upper Sorbian are separate languages as somewhat unresolved (Sussex and Cubberley 2011, p. 94).

This view is complicated by the fact that several leading Sorbian scholars regard Sorbian as a single language with two literary variants, Upper Sorbian and Lower Sorbian. The argument that Sorbian is a single language seems to have political overtones – if it is a single tongue, then the number of its speakers will be inflated (Toops 2001).

One argument is that the existence of transitional dialects means that Sorbian is a single language existing in a dialect continuum (Moseley 2010). Yet the Czech-Slovak and Ukrainian-Belarussian language pairs also have transitional dialects, and those are regarded as four separate languages, not two languages. Further, when Sorbs use the word *Serbski* (*Sorbian*), they refer to either Upper Sorbian or Lower Sorbian but never to both languages together (Toops 2001).

Upper Sorbian has many more speakers than Lower Sorbian and is

still spoken by children. It is expected to survive the 21st Century. Upper Sorbian has 20,000 speakers and in many places, they have their own Upper Sorbian schools. It is spoken in Upper Lusatia around Bautzen. But schools have been closing in recent years. Reasons are controversial but the best analysis seems to be that birth rates are dropping across Germany. All speakers are bilingual and more than half of them use more German than Upper Sorbian. Upper Sorbian shares a lot of unique vocabulary with Czech that is not found in the rest of Slavic.

Intelligibility estimates other than found in formal intelligibility testing were hampered in this study due to a lack of informants from either Upper or Lower Sorbian.

Upper Sorbian intelligibility of other Slavic lects:

Intelligibility figures for Upper Sorbian was hampered by a complete lack of informants.

Upper Sorbian has very high intelligibility of 82% of the extinct Polabian language which was formerly spoken in the area.

Upper Sorbian intelligibility of Lower Sorbian has been [estimated at 80%](#). However, that figure is probably no more than an educated guess. Although some reports say that Upper and Lower Sorbian are [mutually intelligible](#), expert opinion says otherwise. Intelligibility between the two languages, though high (Moseley 2010), is not full. Upper Sorbians do not understand Lower Sorbian well without some formal study of the lexical and phonological differences between the two (Toops 2001).

Upper Sorbian has 33% intelligibility of Slovak, with a range of 26-39%* (Sloboda and Brankaiekc 2012).

Upper Sorbian has 31% intelligibility of Czech, with a range of range 20-40%* (Sloboda and Brankaiekc 2012).

Upper Sorbian has 8% intelligibility of Polish, with a range of 0-18%* (Sloboda and Brankaiekc 2012).

Upper Sorbian has been the subject of an intelligibility study. This study tested a group of L1 Upper Sorbian speakers in high school on their knowledge of Czech, Slovak and Polish. They understood Slovak best, then Czech, and finally Polish. The three results are given above.

Lower Sorbian has fewer speakers at 5,000. It is spoken in the

lower part of Lusatia around Cottbus. There are still a lot of Lower Sorbian cultural activities in Cottbus – a Sorbian kindergarten, school holidays and festivals, etc. It sounds more Polish than Upper Sorbian does. There are still some children learning the language, but most of the speakers are older. Lower Sorbian is predicted to go extinct in this century.

Lower Sorbian intelligibility of other Slavic lects:

There were no Lower Sorbian informants, which made MI study most difficult.

Lower Sorbian intelligibility of Upper Sorbian is estimated at 80%, but that is probably little more than a guess.

Polish

Polish is actually a macrolanguage consisting of four and possibly eight languages: Polish, Kashubian, Silesian, and Cieszyn Silesian are surely separate languages, and North Kashubian, South Kashubian, Opole Silesian, Gilwice Silesian, Podhale, and Mazovian could possibly be seen as separate languages on MI grounds.

There are said to be dialect chains connecting Polish with Belarussian and West Ukrainian. The Rusyn dialect Lemko could be seen as Polish-West Ukrainian transitional. West Ukrainian itself could be seen as Polish-Belarussian transitional.

Most dialects of Polish are mutually intelligible. The differences in Polish dialects used to be more marked, but in recent years there has been a lot of convergence and leveling, possibly due to effects of World War 2.

A good argument supporting the idea that Silesian and Kashubian are separate languages from Polish is that both have undertaken independent development for at least the last 700 years. The influence of either language on Polish has been negligible, although Polish itself has had major effects on both languages.

Instead, Standard Polish is mainly made up of the Małopolski or Lesser Polish and the *Wielkopolski* or *Greater Polish* dialects, with some input from the *Mazovian* dialects.

Mazovian may actually be a separate language, as Polish has [difficult intelligibility](#) with Mazovian. Mazovian is one of the largest and most diverse Polish dialects. It is spoken in the northeast. Until 1918, it was actually a separate language from Polish called the Mazovian language, standardized and codified. After that date, the Modern Polish was created, ending the Mazovian language.

Polish was created by combining a large amount of the Wielkopolski or Greater Polish dialects with a smaller amount of Mazovian. It is hard to say when Mazovian split off from the rest of Polish, but Mazovia was a separate Duchy from the 1100's-1500's, so it was definitely separated from Polish during that time. After that time it was reunited with Poland. If Mazovia and Polish have been separated for 500-900 years, that length of separation typically produces 25-55% MI. But Mazovian came back under the influence of Polish in the 1500's, so the figure may be higher than that.

Polish intelligibility of other Slavic lects:

Polish MI studies were severely hampered by the lack of Rusyn, Lemko, Podhale, Western Ukrainian, and Mazovian informants. However, for Polish itself there were quite a few Polish informants, including one slightly helpful personal informant.

Polish ethnic nationalists are very common, and Polish linguistics has long had nationalist overtones. This nationalism has not died out since the break with Communism, and in fact it has gotten a lot worse. Ethnic nationalism is worse in Poland than in most of the rest of East Europe outside of the Balkans, and it shows no signs of going away. In fact, Poles still see themselves in the midst of a nation-building project. Much rewriting of history and changing fact to nationalist myth has accompanied this endeavor.

The history of the World War 2 era in particular has been badly manipulated in Polish schoolbooks. Hatred of Russians is probably greater in Poland than anywhere else in Eastern Europe, along the lines of the rabid Russiaphobia seen in the Baltics since Independence. Poles never took to Communism well (Stalin once said that imposing Communism on the Poles was like putting a saddle on a cow), possibly due to the heavy influence of the Catholic Church.

Polish ethnic nationalists claim that all Slavic lects spoken in Poland are dialects of Polish and that the notion of other Slavic lects in Poland

was invented by foreigners to “divide the Polish people,” a classic ultranationalist line. They state that Kashubian, the Silesian languages, Podhale, Lemko Rusyn, and Polachian all have full MI with Polish and therefore are all dialects of Polish. This is not true. Even Polish linguists have taken this line for decades, to their discredit.

Any judgment that seemed to be coming from a Polish ethnic nationalist was discarded as suspect.

Poles have 94% intelligibility of Silesian Polish, a dialect of the Polish language. This is just a guess but it is probably about right. Much of what is called Silesian nowadays is simply a much watered down version that is mostly Polish with some German words. Polish-Silesian intelligibility is hard to study due to the confusion and conflation between what is the actual Silesian language and what is simply this Silesian dialect of Polish.

Polish intelligibility of Lemko Rusyn has been estimated at 77%. That is just a rough guess, and the real figure could be a ways off from that. Polish intelligibility of Rusyn is not known, but Poles can understand Lemko Rusyn, the type spoken in Poland, much better than Ukrainians can. This implies that Lemko Rusyn is heavily-Polonized.

Polish understanding of Rusyn has been estimated at 75%, but this is mostly just a guess.

Polish comprehension of Slovio, the Slavic constructed interlanguage, is 70%. This is pretty good intelligibility. One wonders how the rest of the Slavic languages score on this constructed language.

Polish understanding of Eastern Slovak is estimated at 70%, but this is just a guess.

Polish intelligibility of Podhale is estimated at 63%. It is quite clear that most Poles cannot understand Podhale at all. Reports saying that Poles cannot even get the gist of a conversation imply quite low MI for at least that speaker. Podhale has been separated from Polish for a good 400 years.

Based on a rough measure of glottochronology used to estimate intelligibility based on years of separation, an estimate of 63% intelligibility was computed. Ukrainian/Russian have been separated 500 years and have 50% intelligibility. Belarussian has been separated from Russian and Ukrainian for ~300 years and has 75-

80% intelligibility of each. Therefore, 400 years' separation would be between 77% and 50%, which is 63%.

Poles understanding of Western Ukrainian is 46%.

Polish intelligibility of spoken Slovak is a bit higher than that of Czech, 41%* (Golubovic and Gooskens 2015a). Written is quite good at 54%* (Golubovic and Gooskens 2015a).

Polish understanding of Kashubian is less than 41% (Majewicz 1986). Reports that Kashubian and Polish are mutually intelligible are in error. They are motivated more by politics than linguistics (Hopkins 1996). Poles report that they can read Kashubian well, and sometimes they can understand slow, careful speech. But when Kashubian is spoken at a normal pace, sometimes they cannot pick up anything at all.

Intelligibility testing revealed that Poles understand Kashubian worse than they understand Slovak* (Majewicz 1986). However, intelligibility figures in percentages were not available from this study.

Since Poles have 41% intelligibility of Slovak (Golubovic and Gooskens 2015a), a good analysis of Majewicz' study would mean that Poles have less than 41% understanding of Kashubian.

Polish has difficult intelligibility of spoken Czech at 36%* (Golubovic and Gooskens 2015a). Poles understand 50% of written Czech* (Golubovic and Gooskens 2015a). Czech is hard for Poles to understand because Czech has quite a few words that are shared only by Czech and Upper Sorbian and not by any other Slavic languages. At one time, there was probably a dialect continuum between Czech and Upper Sorbian, but Germans moved into the area, breaking the continuum. On the other hand, Czech is fairly easy to pick up for them with concerted study.

Polish oral intelligibility of Belarussian is 32%, and written is 65-70% (Mezentseva 2014). Polish speakers can understand Belarussians about as well as vice versa.

Polish understanding of Standard Ukrainian is lower than of Western Ukrainian at 30%. Written is the same as Belarussian at 67%. The written figure is probably quite accurate.

Polish intelligibility of Silesian is 26% if they have no knowledge of German or Old Polish. Polish lacks full intelligibility of Silesian,

although this is controversial. Some Poles say they find Silesian harder to understand than Belarussian or Slovak. This would line up with our figure above.

Figures for Polish understanding of Silesian are all over the place. In addition, many Poles state that they have full intelligibility of Silesian. This sounds dubious as Silesian lacks full intelligibility of itself. If Silesian cannot understand itself, how can Poles understand it? The argument is irrational.

One problem is that there seems to be a watered-down Silesian that is now little more than a Polish dialect with some German words. This may well be easily understandable by Poles. However, the hard, more pure Silesian is certainly not fully intelligible with Polish.

In addition, there seems to be a lot of individual variation in Poles' ability to understand Silesian. One factor is intelligence and knowledge. Poles who have some knowledge of German and Old Polish understand Silesian much better than those who do not.

Worst of all, the issue has become wildly politicized, has taken an ugly and confrontational tone, and has been dominated by high emotionalism and nationalism. There is a high probability of Polish nationalist bias throwing intelligibility figures off.

Informal intelligibility testing showed that Poles had only 26% intelligibility of Silesian if they have no knowledge of German or Old Polish. With such knowledge, the figure can go up quite a bit.

Polish has low intelligibility of spoken Serbo-Croatian at 19%* (Golubovic and Gooskens 2015a). Written is higher at 33%* (Golubovic and Gooskens 2015a). Polish has poor intelligibility of all of South Slavic, similar to Czech, another major West Slavic language.

Poles understand little of Slovene at 15%* (Golubovic and Gooskens 2015a). Written is a bit better at 26%* (Golubovic and Gooskens 2015a).

Poles have very low intelligibility of only 14% of Upper Sorbian and 12% of Lower Sorbian, however, some Poles score much higher. However, the 14% lines up well with Silesian-Polish intelligibility, so it looks good in MI terms.

Poles understand spoken Bulgarian poorly at 13%* (Golubovic and Gooskens 2015a). Understanding of written Bulgarian is much better,

36%* (Golubovic and Gooskens 2015a). The scenario of very low spoken understanding and much better written comprehension seems to be playing out here too. Written Bulgarian is similar to written Russian as both were based on Old Church Slavonic, and Poles understand written Russian very well, so this may be the reason for the relatively high written MI.

Polish intelligibility of Macedonian is the same as for Bulgarian at 13%. This figure looks good as it lines up with scientific intelligibility testing of Bulgarian.

Poles understand Lower Sorbian slightly worse than Upper Sorbian at 12%.

Polish has low intelligibility of Russian at 7% (Kmiecik 2014). However, Poles understand written Russian very well at 70%. This is a similar situation for Czech, which also had very low oral MI and very high written MI. I am not sure what the reason for the discrepancy is.

Polish speakers state that communication with Russians is impossible. As with so many Slavic language pairs, the presence of false friends between Polish and Russian further complicates matters for spoken understanding. For instance, Russian *зовут* or *zavyud* "they call", sounds similar to Polish *zawód* "a job". Russian *отдыхать* or *odichatch* means "to rest or relax", whereas Polish *oddychać* means "to breathe" (Kmiecik 2014).

Podhale or Goralski

Podhale or Goralski is a Polish dialect that is actually a separate language spoken by the Goral highlanders of the Tatra Mountains that has difficult intelligibility for some Polish speakers, especially those from the cities and the northwest. Podhale is part of the *Małopolski or Lesser Polish dialect* system.

The Lesser Polish *Spisz dialect* is spoken in Spisz and Orawa in the same region by Slovaks who live in Poland and is very close to Podhale. The Spisz dialect is better seen as part of Podhale (Hannan 1996b). Across the border in Slovakia, Podhale is spoken by some Poles in the cities of Sucha Hora and Zdiar. Traditionally, residents of

the area were multilingual, speaking Polish, Slovak and even German and/Hungarian.

Podhale has significant influence from the Romance language Vlach, which is close to Romanian.

This influence comes from Vlach or Wallachian shepherds who migrated to the Tatras from the 1300's-1600's. Lemko Rusyn, with which Podhale has so much in common, also has a lot of Vlach influence from the same source (Lund 1997).

Podhale is composed of a mixture of Polish and Vlach with a significant Slovak layer superimposed on it.

Podhale also has a number of "Carpathianisms" or words peculiar to the general Carpathian Region (Hannan 1996b). Presently Podhale is under very heavy influence of the Polish language.

There are 11 different Goral groups, and each one speaks a different dialect. The most different one is *Silesian Podhale* which has some words from the neighboring Silesian language. Intelligibility is good between all 11 dialects.

Podhale is well understood in nearby Krakow, where another Lesser Polish dialect, *Krakow Polish*, is spoken. However Poles living further away have a difficult time understanding Podhale. In Northwestern Poland, intelligibility with Podhale is so difficult that speakers from this region cannot even get a general idea of what the topic of the conversation is. Some Poles report that Podhale is incomprehensible to them.

Poles visiting the area of Zakopan in the Tatras say that the Podhale-speaking Goral Slovaks from across the border seem as if they are speaking Polish until they listen closely and realize it is another language altogether. That they think it is a separate language implies that they do not think it is a Polish dialect and also that MI might be difficult.

The truth is that most Poles cannot understand Podhale. However, Podhale is more intelligible to Poles than Kashubian.

Podhale has a lot in common with Cieszyn Silesian and especially the Lemko Rusyn spoken by the many Lemkos who live nearby. Podhale's relationship and MI with Cieszyn Silesian and especially Lemko Rusyn remain uncertain. In addition, Podhale may have a close relationship

with the nearby Eastern Slovak language, but just how close it is to Eastern Slovak is not yet known, and MI is uncertain also.

Podhale seems to be part of a "language area" consisting of Podhale, Slovak (especially Eastern Slovak), Rusyn (especially Lemko), the nearby Lesser Polish dialects, and Silesian (especially Cieszyn Silesian). The precise relationships and MI between these various lects has not been fully explored.

Poles in general have difficult intelligibility of Podhale. A good case can be made that Podhale is a separate language on the basis of mutual intelligibility.

Kashubian

Kashubian may be a macrolanguage made up three different languages different languages, [North Kashubian](#), [South Kashubian](#) and Central or Standard Kashubian.

North and South Kashubian have difficult intelligibility (Martin 2013, Hopkins 1996).

Kashubian was the same language as Polish when Polish began forming in the 1100's (Sussex and Cubberley 2011, p. 91). But in the 1300's, it started to break away and become a separate language (Martin 2013). It is spoken in the far north of Poland around the Gdansk, Sopot, and Gdynia area.

Polish linguists have long held that Kashubian is a dialect of Polish, and a large number of Polish linguists hold this opinion to this very day.

Although they claim to be making this determination on pure linguistic grounds without any political bias, this claim is suspect, as the status of Kashubian as a language or dialect has been a hot-button political issue in Polish politics since Poland was first established as a nation-state after World War 1 (Hopkins 1996).

Poland's access to the sea at the time depended on the Kashubian area being populated by Poles (Hopkins 1996).

After World War 2, any discussion that Kashubian was a separate

language was equated with an effort by Germany to undermine Poland's right to possession of the Kashubian corridor and second by a desire by Kashubians (who are often seen as unpatriotic) to establish a separatist movement and break up the geographical integrity of Poland. In truth, there is no Kashubian separatist movement (Hopkins 1996).

Hence, discussion of the Kashubian Question was banned under Communism. The modern Kashubian movement is not separatist, but it does oppose Polish ethnic nationalism and promotes the notion of a multi-ethnic state. Polish nationalists oppose this idea and continue to wave the non-issue of Kashubian separatism to bolster their opposition to a more diverse Poland (Hopkins 1996).

Kashubian has an uncertain number of speakers. Estimates run the gamut from very high to very low figures. At the high end, it is said that Kashubian has 200,000 speakers.

The low estimate is 3,000 speakers (Sussex and Cubberley 2011, p. 91).

Both figures appear to be incorrect. There may have been 200,000 speakers 50 years ago, but there are not nearly that many now. The Polish census figures were distorted because Kashubians were not allowed to choose Kashubian as a nationality, so many checked Kashubian language instead as a way of affirming their ethnic heritage.

3,000 speakers is certainly much too low, as Kashubian is taught in 80 schools and is a co-official language in 10 towns that have more than 20% Kashubian population (Martin 2013).

A much better figure is 53,000 speakers (Martin 2013).

Ethnologue lists Kashubian as moribund, but this is probably not correct, as moribund means the language has only elderly speakers. Young Kashubian speakers exist, and the language is taught in quite a few schools, although the number of youthful speakers is quite low. Kashubian is also spoken in Canada, where there are now 7-8 generations who have spoken the language. However, the youngest speakers are in their 50's, and Canadian Kashubian organizations estimate that the language will only survive 30 more years in Canada.

There are no intelligibility figures for Kashubian, but figures for North and South Kashubian were obtained. However, they were basically no

more than educated guesses. Nevertheless, the real figures are probably not far off from this guesswork.

North Kashubian or Północnokaszubskiego Kashubian is spoken in the far north along the coast of the Baltic Sea and is the furthest Kashubian dialect from Polish. Many North Kashubian speakers were considered so Germanized that most of them were resettled in Germany after World War 2. North Kashubian is now spoken in only a few villages due to massive depopulation. North Kashubian had a close relationship with the Slovincian language which went extinct in the mid-1900's, with traces remaining alive until the 1960's in one town.

MI research with Kashubian was severely hampered by a complete lack of Kashubian informants.

North Kashubian intelligibility of other Slavic lects:

North Kashubian has 85% intelligibility of the extinct Pomeranian language.

North Kashubian has 85% intelligibility of the extinct Slovincian language.

North Kashubian has high but not full intelligibility of South Kashubian at 80%.

Southern Kashubian or Południowokaszubski Kashubian is the Kashubian dialect that is closest to Polish.

South Kashubian intelligibility of other Slavic lects:

South Kashubian has marginal understanding of the extinct Pomeranian language at 85%.

South Kashubian has very high but not full understanding of Slovincian, an extinct language.

South Kashubian has marginal understanding of North Kashubian at 80%.

Silesian

Silesian or Upper Silesian, possibly a macrolanguage, consists of Silesian and possibly two more languages - Opole Silesian and Gliwice Silesian. Silesian itself consists of Opava Silesian and Jabłonków

Silesian, both spoken in Czechia, and Niemodlin Silesian, *Kluczbork Silesian*, *Prudnik Silesian*, *Opole Silesian*, *Gliwice Silesian*, and *Sulkovian Silesian*, all spoken in Poland (Martin 2013). *Niemodlin Silesian* is Silesian-Polish transitional.

Silesian or Upper Silesian, spoken in Poland and Czechia, is the standard language and has the most speakers. Cieszyn Silesian, spoken in Czechia and Poland, is often said to be a dialect of Silesian, but it is surely a separate language. Opole Silesian and Gliwice Silesian may be considered separate languages on MI grounds.

There are many Silesian speakers who do not live in the Opole or Katowice regions, and they also speak Silesian, so we would need a Standard Silesian to account for them.

In Czechia, Opava Silesian and Jabłonków Silesian are [nearly extinct](#) with the exception of a few Moravians who speak these dialects along the Polish border near Racibórz.

After WW2, most speakers of Silesian in Czechoslovakia were said to be Germans and [were deported](#) to Germany. From 1945-1948, over 1 million Germans were killed as Germans were ethnically cleansed from much of Eastern Europe. In Silesia, even local Jews and Poles were killed during the anti-German campaign. Many Poles, particularly from former Eastern Poland (now parts of Belarus and Ukraine) were put on trains and moved to Silesia to replace the Germans who had been ethnically cleansed.

78% of the prewar Silesian population fled to Israel or the US or were expelled to Germany. The ethnic cleansing was most complete in Lower Silesia where a German language, Lower Silesian, was spoken. Speakers were mostly deported to Germany after World War 2 and as a consequence, the language is now nearly extinct.

In fact, most Silesians identified as Poles, not Germans. A study by the British Secret Service found that 92% of rural Silesians were both anti-German and anti-Communist. In fact, they had been heavily persecuted under Nazi rule as anti-Nazis, and there were a number of massacres of Silesians under Nazi rule. When the Communists took over, they were persecuted again, this time as anti-Communists.

In 1977, Silesians were forbidden to give their children Silesian names and had to give their offspring Polish names instead.

The Polish attitude towards Silesians makes no sense. First, Silesians

are seen as Germans and not Poles, hence they are traitors. At the same time, the Silesian ethnicity is said to not exist, as all Silesians are simply Poles. So Poles view Silesians as Poles when it is convenient to do so and as German aliens when that is convenient.

The Poles have recently arrived at a clever compromise which enables them to have their cake and eat it too.

Silesian is regarded as an "ethnolect," that is, Silesians are not Poles (satisfying Polish nationalists who see them as quasi-German aliens), yet they speak a Polish dialect (satisfying Polish nationalists who wish to deny that any languages other than Polish are spoken in Poland).

Silesian in general consists of many dialects with deep divergence between them. Attempts have been made in recent years to create an artificial Silesian koine that can be understood by all speakers, but the Silesian lects are so far apart that [hardly any Silesians can understand this new language](#). There have been numerous attempts to create orthographies for Silesian. Presently in the Silesian Wikipedia, two different orthographies are used.

Ślabikōrzowy Szrajbōnek is the latest Silesian orthography, and it has been approved by the Pro Loquela Silesiana organization as a unified orthography that reflects the sounds of all Silesian lects (Syniawa 2010). Part of the problem is that all of the writing systems are trying to accommodate all Silesian lects in a single system. Silesian is sometimes written in the Polish alphabet, but this is a dubious endeavor as it does not allow for the representation of all of the sounds of Silesian.

Another orthography used is based on Steuer's Silesian Alphabet, created by Felix Steuer in the Interwar Period as a system for writing his poems. In 2006, the Silesian Phonetic Alphabet was created based on Steuer's orthography. Although it is the most phonologically accurate Silesian orthography, it did not become popular with major Silesian organizations because it contained too many caron symbols (ˇ) and hence resembled the Czech alphabet too much and they wanted something more distinct.

Silesian is spoken in Poland and Czechoslovakia and is often thought to be halfway between Polish and Czech, similar to Sorbian. It is spoken in Southern Poland near the Czech border and over the border into the northern parts of Czechia. Silesian is thought to be derived

from Old Polish.

It has been split from Polish since the late 1200's, as the first document to contain a sentence written in the Silesian language (or actually, Old Silesian), *The Book of Henryków*, was written in 1270. This means that Silesian has been split from Polish even longer than Kashubian has, as Kashubian started to break away 75 years later. During this long period, Silesian underwent heavy German influence.

There are different theories about the history and genesis of Silesian. Perhaps the best theory is one that is not written about much. It goes something like this:

The ultimate base of the Silesian language (called *Shluenski* in the region or *Slůnsko* in Polish) is Celtic. In 400 BC, all of the towns and cities in Silesia had Celtic names. Silesia was invaded by Romans 2,000 YBP. Some time later, Ashkenazi Jews moved into Silesia in large numbers. In fact, the Silesian and Sorbian areas in eastern Germany and western Poland are thought to be the ultimate homeland of the Yiddish language. For some time, the people of the region were Celtic-Ashkenazi Jewish. Silesians were regarded as free tribes during this era from the 600's to the 800's-900's

The people of Silesia at this time were a mix of Celts, Ashkenazi Jews, Jors (Bohemian Celts), and Lombards.

From the 800's to the 900's, Silesia was part of the Kingdoms of Moravia and Bohemia. During this period, highlighted by the reign of King Barbarossa, the Silesians were speakers of a Celtic language.

The Lombards were a Germanic tribe that later moved into northern Germany in large numbers. The Lombards arrived in the area before the Proto-Slavic and Germanic speakers.

About 900 YBP, Proto-Slavic and Germanic-speaking tribes moved into the area.

Texts in Old Silesian show heavy Germanic and Yiddish influence.

From the 900's to the 1200's, Silesia was a part of Poland. It would not be a part of Poland again until 1920.

At some point later, after 1100 CE, a significant number of Gypsies or Roma moved into the area.

From 1300-on, Silesia was alternately independent and was ruled by Czechs, Austrians and Germans.

In 1335, Silesia came under Czech rule (Hannan 1996b).

From the late 1400's to the early 1500's, Silesia was ruled by the Austro-Hungarian Empire. From the 1400's-1500's on, the principal language in Silesia was either German or Czech (Hannan 1996b).

In the Middle Ages, while Silesia was part of the Austro-Hungarian Empire, there over a dozen different semi-autonomous Duchies or principalities in Silesia at various points in time (Hannan 1996b).

In the 1700's, most of Silesia was ruled by Prussia, but a small part of it was ruled by Austria (Hannan 1996b).

Much later, others moved into the area. French moved in with Napoleon's army. After Napoleon was defeated, English settlers moved in. At some point, a significant number of Swedes settled in Silesia.

Silesia switched back and forth between German and Polish rule. During periods of Polish rule, the area would be Polonized. During periods of Germanic rule, Germanization occurred. When Napoleon moved in, there were attempts to impose French on the population.

From 1920-1939, 30% of Silesia was an autonomous province in Poland, while the other 70% of Silesia was in Germany. After 1945, most of Silesia was in Poland with lesser parts in Czechoslovakia and in Lusatia in Germany. Silesia was only part of Poland for 30% of its history.

Presently, the Polish language is dominant in Silesia, but prior to 1937, that was not the case. The Neisze and Odra Rivers form the northern and southern borders of Upper (Southern) and Lower (Northern) Silesia. As one moves further to the west, the language becomes increasingly Germanic and as one moved east, the language becomes more Polish.

In the center is the more pure language with its Celtic-Yiddish base with areas where Moravian, Czech, Roma, or Yiddish influences are greater. More typically, the language of the center is a combination of all or most of these influences.

There was a French-speaking enclave around Gracze and Nowy Chapel and an English-speaking enclave from Wellington's troops who defeated Napoleon. These soldiers married Polish, Jewish, and German women and remained in the area. French and English

influences remain in certain Silesian lects, but many of the words have been changed under the influence of heavily-Polonized Silesian. During various periods, Danish, Hungarian, French, Swedish, and English farmers moved in to cultivate the rich soil.

In Upper Silesia, the population was about 25% Germanic, 25% Slavic and 25-33% core Yiddish-Celtic.

In Northwestern Lower Silesia, Sorbic, and Upper Saxon influences account for ~47% of the population. The Sorbic group is not made up of Sorbians as one might assume but more accurately consists of speakers of the Gothic, Gotlandic or Gutnish language of Southern Sweden.

In Northeastern Lower Silesia, the same two influences represent over 50% of the population. In Western and Central Upper Silesia, the population is referred to as Bohemian. This population is more accurately a mix of Jor Celtic, Germanic, and Celtic with a lesser mixture of Czechs and Moravians. Later, Slovaks added to the mix. Throughout Upper Silesia there are islands of Polish and German speakers. A core Celtic-Yiddish is mixed all through Upper Silesia but predominates in Central Upper Silesia.

Germanic influence via Upper Saxon is dominant in Lower Silesia. In Upper Silesia, Germanic does not dominate as Polish influence is stronger. The Germanic influence in the south comes more from Bavarian and Franconian German. Later, Gallo speakers from Normandy moved in to Upper Silesia. Gallo is a langue d'oïl spoken in Normandy. It has significant Celtic influence from Breton.

The language itself is at least 30% Germanic. There is also heavy Yiddish influence with many words from Polish, German and other languages combined with Yiddish roots, words, or parts of words to create new words.

The most notable thing about Silesian is the way it is written. Many Silesian words are German words that are written with Polish phonotactics. For instance, Silesian *aszynbecher* from German *Aschenbecher* "ashtray" (Kmieciak 2014).

Silesian has strong German influences. In fact, some Poles, when listening to audio of Silesian speech, claim that Silesians are speaking German. Aware of the Silesians' strong residual German identity, many Poles say they are Germans and not true Poles.

The more German the Silesian dialect is, the harder it is for Poles to understand. In recent years, many of the German words are falling out of use and being replaced by Polish words, especially by young people.

Poles who know German and Old Polish can understand pure Silesian quite well due to the Germanisms and the presence of many older Polish words (Kmieciak 2014), but Poles who speak only Polish have a hard time with pure Silesian.

Many Poles insist that Silesian is a Polish dialect, but this is based more on politics than reality. Linguists have caved in badly to pressures from Polish nationalists over the last 100 years to deny language status to both Silesian and Kashubian. Even in the case of Kashubian, an obvious language if there ever is one, Slavic linguists regard as somewhere in the no man's land between a dialect and a language.

Sadly, Slavic linguists are unanimous that Silesian is a dialect of Polish. This judgment seems quite wrong. At the same time, Slavic linguists are now pressuring the community to regard the fake languages of Serbian, Croatian and Bosnian as full-fledged languages, just because some ethnic nationalists say so. This unfortunate situation seems to reflect an environment where Slavic linguists are captive to political pressures and are all too willing to have politics trump linguistic science in terms of the language-dialect question in Slavic.

Looking elsewhere in Europe, clearly if Galician, Limburgish, Walloon, and Scots are separate languages, as in fact they are, then by the same criteria used to classify these lects, clearly Silesian is a language also. The simple truth is that it is not possible to say on the one hand that Sicilian, Galician, Limburgish, and Scots are languages on the one hand while saying that Silesian is not on the other hand.

Silesian has been issued an ISO code, which means that the formal scientific organization of linguistic science in charge of labeling lects as either dialects or languages has decided that Silesian is indeed a language. In other words, linguists have given a formal stamp to Silesian's status as a language via their official scientific organ in charge of the classification of lects.

Finally, the Polish government itself has recognized Silesian as an

official regional language. Not a regional dialect, but a regional language. It seems odd that the Polish government says Silesian is a language while almost all Poles and Slavic linguists continue to insist it is but a Polish dialect. The argument that Silesian is a dialect of Polish does not make sense on many levels, and this is only one of them.

The truth is that Silesian is a Polish dialect, but it is a dialect of Old Polish rather than Modern Polish (Martin 2013). In fact, people in the Silesian area regard Silesian as incomprehensible. Opava Silesian in Czechia is also incomprehensible to Poles. 40% of Silesian vocabulary is different from Polish, mostly Germanisms. The German influence is more prominent in the west; Polish influence is greater in the east.

Many Silesian speakers now speak a watered-down version of Silesian which is more properly seen as a Polish dialect with some Silesian words. Pure Silesian appears to be a dying language (Kmieciak 2014).

Cieszyn Silesian and Lach Silesian are listed as dialects of Silesian spoken in Czechoslovakia, but the truth is that these are completely separate languages.

Probably at one point they were dialects of Silesian along the Silesian dialect continuum. The continuum probably looked something like Polish-Silesian-Cieszyn Silesian-Lach-Czech. The entire Silesian dialect continuum – Silesian-Cieszyn Silesian-Lach – is Polish-Czech transitional (Hannan 1996b).

But this continuum has been breaking up since 1900, as the Silesian lects on the Polish side have been getting closer to Polish, and the ones on the Czech side have been getting closer to Czech, hence there is a sharp break at the border of the former continuum. At the same time, Cieszyn Silesian and Lach have separated further from Silesian and become separate languages (Hannan 1996a).

An example of the changes is that Lach in Czechoslovakia used to be considered Polish-Czech transitional, but now there is a sharp break at the border and Lach is clearly part of Czech (Hannan 1996b).

Silesian intelligibility of other Slavic lects:

There were serious problems studying the MI of Silesian. The first was an utter lack of Silesian informants. The second was the distortion that Polish nationalists introduced into the debate. Many of their judgments had to be rejected.

Silesian intelligibility of Polish or Czech is hard to measure because all Silesian speakers speak either Polish or Czech. One would have to find Silesian speakers who speak either only Polish to test Silesian intelligibility of Czech or those who speak only Czech to test Silesian intelligibility of Polish.

Silesian lacks full intelligibility of Cieszyn Silesian, as speakers of the two languages communicate with considerable difficulty. A good figure is not available.

Opole Silesian, spoken in Poland near the Czech border, is not intelligible with Katowice Silesian to the northwest, so it is a separate language.

Linguists have noted that Opole Silesian has much less German in it than other Silesian lects (Stadniczenko 1947-48).

Katowice or Gliwice Silesian, spoken in Poland to the northwest of the Opole Silesian border language, cannot be understood by Opole Silesian speakers, so it should be a considered separate language.

No figures for Katowice Silesian - Opole Silesian MI were available.

Cieszyn Silesian

Cieszyn or Těšín Silesian, Ponaszymu/Po Naszymu, or Po Našymu is a language closely related to Silesian spoken by ethnic Poles in Czechia in the far northeast of the country near the Polish and Slovak borders and to a much lesser extent across the border in Poland. The term Po Našymu means "in our own way" (Hannan 1996b).

Some of these Poles call themselves Poles, and others call themselves Czechs depending on which country they live in. It has strong German influences like Silesian but differs from the rest of Silesian in that it has undergone heavy Czech influence.

Some say it is a part of Czech, but this is not correct. Czech linguists say it is Polish-Czech transitional, and Polish linguists say it is a Polish dialect. It is best characterized as Polish transitional to Czech.

This was formerly a dialect of Silesian, but it underwent so much

Czech influence that it drifted apart from Silesian Proper and is no longer intelligible with it.

It was much more Polish prior to 1930, but since then a lot of Czech has gone in.

The Czech influence has been particularly heavy in the past 30 years (Hannan 1996a).

Cieszyn Silesian continues to be spoken by all age groups and is still being learned by children. Cieszyn Silesian is doing better in this region among the local Poles. The Poles in this region are trilingual, speaking Cieszyn Silesian, Czech and Polish fluently. It is not doing as well among the local Czechs, as they are prone to assimilation to the Czech language. In addition, there are also fellow Czechs around them who do not speak Cieszyn Silesian, having moved there from other parts of Czechia.

Although there are a few speakers in Poland, Cieszyn Silesian is not doing as well there, as the speakers are susceptible to assimilation to the Polish language.

Cieszyn Silesian intelligibility of other Slavic lects:

A dearth of Cieszyn Silesian informants made MI studies with this language difficult.

People observing conversations between Cieszyn Silesian and Upper Silesian speakers report that speakers of these languages have a hard time understanding each other, but intelligibility figures are not known. Cieszyn Silesian speakers strongly reject the notion that they speak the same language as Upper Silesians.

Cieszyn Silesian has a close relationship to Goralski Polish or Podhale, which is spoken by the Gorals in the Tatra Mountains.

It is also close to Lemko Rusyn, the heavily-Polish Rusyn that is spoken in the mountains near the Gorals.

Cieszyn Silesian MI with Podhale and Lemko is not known.

Cieszyn Silesian-Lach intelligibility is surely not full, but no figures are available. The two languages are close, both being Polish-Czech transitional.

Cieszyn Silesian understanding of Polish is hard to figure because almost all speakers also speak Polish.

Cieszyn Silesian intelligibility of Czech may be 35%. Ponaszymu appears to lack full intelligibility with Czech. In fact, in some cases, the MI between the two is zero. Cieszyn Silesian intelligibility of Czech has been estimated at 70%. This is probably little more than an educated guess. The true figure may be between the two, which would be 35%.

East Slavic

Belarussian

Belorussian is one of the most recent East Slavic languages to come into existence, as the earliest Belarussian texts are from only the 1500's. So the split between Belarussian and Ukrainian and Russian is shallower than that between Spanish and Portuguese. Belarussian, like Ukrainian, has borrowed many Polish words, but the pronunciation is more like Russian.

One theory has it that in a sense there is no Belarussian language. Instead the north of Belarus is the remains of the old Novgorod Russian dialect (probably the best example of the highly controversial North Slavic branch of Slavic) and the south of country is the remains of the Brest dialect of Ukrainian.

While West Polesian is indeed part of this Brest speech area, the Novgorod dialect went extinct 500 years ago. But if you want to say that Belarussian is a hybrid between an ancient Russian northern dialect and an aberrant Ukrainian northern dialect, you would be on firmer grounds. Nevertheless, Belarussian is definitely a full language, having split from both Russian and Ukrainian 300-400 years ago. This temporal degree of split makes sense in light of 75-80% MI between Belarussian and both Russian and Ukrainian.

Part of the problem above is that whatever its origins, West Polesian is no longer a Ukrainian dialect but instead is probably best seen as purely Ukrainian-Belarussian transitional and not part of either language per se.

Reports of the endangerment or looming death of Belarussian are usually politically motivated attacks on President Lukashenko accusing him of killing the language.

On the contrary, Belarussian, while in a disappointing situation, is very much alive. Almost all Belarussians can speak the language, but only 15% do so in day to day conversation. Most of the rest more often play the role of passive speakers, although they can speak the language if they need to (Mezentseva 2014). 90% of Belarussians actually state that Belarussian is their mother tongue.

Belarussian knowledge of their language benefits them because it gives them a head start on learning other Slavic languages (Mezentseva 2014).

Belarus was actually part of Poland at one time, as was Western Ukraine. Belarussians see themselves as a different people from Russians.

For centuries, they called themselves Tutejshiya "our people" (Mezentseva 2015).

Part of the blame for the decline of Belarussian lies with Belarussians themselves because despite the statements in the paragraph above, Belarussians have a very strong attachment to Russia and only a weak attachment to their own land (Mezentseva 2014). The result of this is that although 85% of Belarussians can speak Belarussian, Russian is the preferred language in the country (Pavlenko 2006).

In 1991, Belarus only had one official language, Belarussian, though Russian was in wide use. In 1994, the people voted to have two official languages, Belarussian and Russian. Russian-language media and politicians quickly took advantage of the situation and used the opportunity to make Russian the dominant language in the country (Mezentseva 2014).

Lukashenko regularly wins elections by 75-80% margins, and polls show about the same support. The very unpopular opposition are regarded by most Belarussians as traitors and anti-Russian, pro-US tools of the West out to destroy the country.

One major problem for the language is that Belarussian is now associated with the opposition in the country. This association of the language with the unpopular opposition has hurt the language and is a major reason why state support for Belarussian has been lukewarm

at best (Mezentseva 2014).

However, the linguistic situation in the country is complicated, and there are Belarussian-language TV stations and a number of daily newspapers (Mezentseva 2014).

The Western media reports that Belarussian is dying, but this is politicized discourse.

The truth is that Belarussian is becoming more popular these days, as it is coming to be seen as the prestigious “language of the intelligentsia,” as opposed to the Soviet era in the 1970's and 80's when it was regarded as a “village language.” Belarussian language advocates say that they are not pessimistic at all about the state of the language, and in fact they are optimistic. Belarussian is used in the educational system, and advocates expect its use there to expand. Independent Belarussian classes have been springing up to assist Belarussians who want to promote the language and culture. (Mezentseva 2014).

West Polesian or West Palesian is a transitional Belarussian dialect to Ukrainian. Some think that West Polesian is a microlanguage, but the majority of Belarussian linguists say it is a dialect of Belarussian (Mezentseva 2014). On the other hand, Ukrainians listening to this dialect insist that it is Ukrainian spoken in Belarus. See the section on this dialect under Ukrainian for more on this confusing language.

Since Ukrainians have full intelligibility of this dialect, it seems hard to make the case that it is Belarussian.

Trasianka is Belarussian dialect based on a mix of Russian and Belarussian that arose during the Sovietization of Belarus. It resembles Russian spoken with a Belarussian accent and is spoken mainly by rural dwellers who moved to towns and started to watch a lot of Russian TV. It is also widely spoken in Eastern Belarus near the Russian border (Mezentseva 2014).

There are still many Belarussian mother tongue speakers in the US, including people in their 20's.

Belarussian intelligibility of other Slavic lects:

Belarussian MI research was helped by the presence of an excellent personal informant who was also a linguist.

Belarussian understanding of Ukrainian is high at 80% (Mezentseva

2014).

Belarussian intelligibility with both Ukrainian and Russian is a source of controversy. Belarussian has transitional dialects that are intelligible with dialects of both Russian and Ukrainian.

Belarussian and Ukrainian have 84% lexical similarity.

Pronunciation is also very similar between the two languages. Some of the grammatical categories do differ (Mezentseva 2014).

Russian nationalists often state that Belarussian is a dialect of Russian. However, this judgment is based more on national chauvinism than linguistics, as Russian lacks full intelligibility of Belarussian (Mezentseva 2014).

Testing Belarussian intelligibility of Russian is not realistically possible. The vast number of Belarussians speak Russian, and of those who do not, all or nearly all have at least passive knowledge of Russian. At the moment there are few to no Belarussian monolinguals. If they exist at all, there may be a few elderly female monolinguals in the far west of the country by the Polish border, but it would be difficult to study them (Mezentseva 2015).

Belarussian has many Polish borrowings, hence Belarussian has a fairly high intelligibility of Polish at 41% (Mezentseva 2015). Written intelligibility is higher at 67% (Mezentseva 2015). Although Polish is notorious for being one of the hardest languages in Europe for foreigners to learn, Belarussians can actually learn it fairly easily due to the similarities between the two languages (Mezentseva 2014).

Belarussian has relatively low comprehension of Slovak at 25%.

Podlachian

Podlachian, Podlachian, Polaskiego, or Podlasie is a small language spoken in far Northeastern Poland near the Belarussian border (Maksymiuk 2014) that, on first glance, appears to be Polish-Belarussian transitional. Podlachian is sometimes also seen as a Rusyn dialect. This is best seen as a Ukrainian lect transitional to Belarussian that is full of Polish words. Experts are uncertain whether this is a dialect of Ukrainian or a microlanguage of its own.

As early as 1884, linguists were already describing this lect as a separate language, the Chachłackiego language (Żelechowskij 1884). The term Chachłackiego is pejorative and is rejected by speakers.

Although speakers speak a divergent dialect of Ukrainian with massive Polish influence, few of the speakers identify themselves as Ukrainian. Of 33,000 speakers, only 1,200 identify as Ukrainians, and the other 30,800 say they are Belarussians (Maksymiuk 2014). Podlachian looks a lot like Torlakian in that speakers say that they speak the language of the ethnicity that they identify with.

For instance, Podlachian speakers who see themselves as Ukrainian say they speak Ukrainian, and those who see themselves as Belarussian say they speak Belarussian, etc. (Maksymiuk 2014).

The number of speakers range from 33,000 (Maksymiuk 2014) to 50,000 in Podlachia. There may be another 500,000 speakers in Belarus and Ukraine, but this has not yet been confirmed.

All Podlachian speakers also speak Standard Belarussian, but this seems to be more because they have learned it than because Belarussian and Podlachian are inherently intelligible. Podlachian speakers recognize that Podlachian is different from Standard Belarussian.

However, Podlachian L1 speakers sometimes do not begin to learn Polish and Belarussian until they go to school. The fact that they have to learn Polish and Belarussian in school implies that Podlachian is neither Polish nor Belarussian.

Podlachians say that they do not speak the same language as either Standard Ukrainian or Standard Belarussian. Podlachian has marked differences between it and Belarussian (Maksymiuk 2014).

In 2011, Podlachian speakers petitioned SIL to have Podlachian recognized as a separate language, but SIL rejected the request on unknown grounds. The SIL request did not discuss intelligibility issues but instead focused on Podlachian being an endangered language in need of protection (presumably SIL recognition would help in this regard) and the significant differences between Podlachian and Ukrainian.

Podlachian intelligibility of other Slavic lects:

Podlachian MI was very difficult to study due to the utter lack of

Podlachian informants. Hence the only figure obtained relied on guesswork.

Podlachian and Ukrainian have 82% lexical similarity. In contrast, Ukrainian has 84% lexical similarity of Belarussian, and Ukrainian-Belarussian intelligibility is 78%. The suggestion is that Podlachian and Ukrainian may have 75% intelligibility.

MI data of Podlachian with Belarussian, Polish, or Rusyn is unknown. Podlachian-Belarussian intelligibility is hard to study because all Podlachians also speak Belarussian.

Ukrainian

Ukrainian is actually a macrolanguage consisting of at least three and possibly five languages: Ukrainian, Rusyn, Podlachian, and possibly Western and Eastern Ukrainian.

Ukrainian does not appear as a discrete tongue until the late 1500's (Sussex and Cubberley 2011, p. 85), so one view is that it has only been split from Russian for only 450 years.

However, another view says that Ukrainian and Russian began diverging 50-100 years after the collapse of the Rus in 1170, which would mean that the two languages have been split for 700-800 years.

At any rate, the time depth of the split seems to be approximately the same as between English and Scots. Scots-English MI has been formally tested at 42% in single very limited pilot study. Scots-English MI is similar to the 50% estimates of Ukrainian-Russian MI.

Before that, Ukrainian had been written as a combination of Bulgarianized OCS and Ruthenian, the language of the Rusyns (Kostomarov 2015). With the Union of Lublin in 1569, Ukraine came increasingly under Polish influence, supplanting Ruthenian and OCS. OCS was reinstated in 1581, and Ukraine saw increasing movements towards the Orthodox Church (Sussex and Cubberley 2011, p. 85).

The movement away from Polonization and Catholicism towards Orthodoxy resulted in the beginnings of a nationalist feeling among Ukrainians. In Western Ukraine, a Ukrainian written language began

to appear under the influence of the Reformation and the translation of religious literature into vernaculars across Europe. This is when Ukraine began "looking to the West," and this Westernization resulted in many borrowings from Western Europe (Sussex and Cubberley 2011, p. 85).

Anti-Polish and anti-Catholic sentiment climaxed in the Khmelnitzky Peasant Rebellion of 1648, an anti-feudal uprising that resulted in an incredible liquidation of 1/3 of the population of Greater Poland. This war led to the creation of a Cossak state in Eastern Ukraine under a Russian protectorate. The Cossacks were defeated by Russia in 1709 and Russia annexed Ukraine. The 1700's saw the decline of Polish influence and the rise of Russian influence (Sussex and Cubberley 2011, p. 85).

At the end of the century, Poland was partitioned again, and Russia ended up ruling most of Ukraine. The 1800's saw increasing Russian influence both politically and linguistically (Sussex and Cubberley 2011, p. 86).

Ukraine began to develop a feeling of true nationalism separate from Russia during this century. OCS was gradually replaced, as it was associated with Russia, and the written language moved towards the vernacular, being based on more on eastern dialects from Kiev to Kharkov (Sussex and Cubberley 2011, p. 86). Ukrainian publishing was somewhat curtailed by the Russians, and the center of Ukrainian literary work moved west to Lvov and Galician Ukrainian became the written language for a period (Shevelov 1981).

When Ukrainian publishing opened up again 1905, the Kiev dialect took over from the Western version.

After the October Revolution, Ukraine became the Ukrainian SSR within the USSR. During the 1920's, a written Ukrainian language was codified using lexis from Galicia and phonology from the eastern dialects. This flowering of a literary language crashed in 1930 with Stalin's crackdown on regional nationalisms. Russian became the second official language for Ukraine and turned into the language of business (Sussex and Cubberley 2011, p. 86).

However, after 1945, finally most Ukrainian speakers were in a single country called Ukraine with the exception of three million Ukrainian speakers in Russia. After 1991, Ukraine gained its independence, and

Ukrainian moved into position as the official language of the state, and a campaign of de-Russification of the language began. Presently, Ukrainian has established itself as a full-fledged national language (Sussex and Cubberley 2011, p. 87).

Ukrainian has borrowed many Polish words, but the phonology more resembles Russian.

There are dialects in between Ukrainian and Russian such as the Eastern Polissian and Slobozhan dialects of Ukrainian that are intelligible with both languages. Complicating the picture is the fact that many Ukrainians are bilingual and speak Russian also.

Kuban Russian or Balachka is a Russian-Ukrainian transitional dialect spoken in the Kuban area right over the eastern border of Ukraine. It and Slobozhan Ukrainian are very close to each other.

It is best seen as a Ukrainian dialect spoken in Russia – specifically, it is markedly similar to the Poltavian dialect of Ukrainian spoken in Poltava, Zaporozhye, and Cherkasy in Central Ukraine, where it is often referred to as *Surzhik*. This is because these people originally migrated to to the Kuban from Zaporozhia. 190,000 of them were resettled by Katharine the Great from Zaporozhia to the Kuban. The original Cossacks on the Dnieper River in Central Ukraine who later moved southwest to the Don area were called Cherkasy. It is mostly Ukrainian with a few Russian words.

It has now been shown via genetics that Don Cossacks are half-Belorussian and half-Ukrainian. There is also a link between Don Cossacks and Novgorod Russians of Northern Belarus and neighboring Russia.

However, Balachka is dying out and is now mostly spoken only by a few old people. There are only a few child speakers. Most people in the region now speak Russian with a few Ukrainian words.

Although the standard view is that Balachka is a Ukrainian dialect, some linguists say that it is actually a separate language closely related to Ukrainian. An academic paper has been published making the case for a separate Balachka language. In addition, Balachka language associations believe it is a separate language. The case that it is a separate language is apparently made on structural grounds, but it that is dubious, as dubious as Balachka intelligibility with Ukrainian is full.

Balachka intelligibility of other lects:

Balachka has [full intelligibility of Ukrainian, 100%](#), which suggests that it is really best seen as a dialect of Ukrainian despite this treatment formally putting it in Eastern Ukrainian.

Balachka-Russian intelligibility is 60% because Balachka is basically Ukrainian spoken in Russia.

West Polissian or West Polesian is a Ukrainian dialect spoken in the far northeast on the border with Belarus in the northern part of Volyn Oblast and the northwestern part of Rivne Oblast in Ukraine and in some districts of Brest Oblast along the border in Belarus. When it is in Ukraine, it has Ukrainian grammar and is called a Ukrainian dialect. When it goes into Belarus, it gets Belarussian grammar and is called a Belarussian dialect. This is a Ukrainian-Belarussian transitional dialect. This dialect cannot be called either Belarussian or Ukrainian; instead it is both.

Ukrainians are emphatic that West Polesian is a Ukrainian dialect.

West Polesian poses difficulties in terms of classification. Linguistically, it seems to be a part of both the Ukrainian and Belarussian languages, which doesn't make sense. It is either a bidirectional transitional dialect of two different languages (an odd way to characterize a dialect), or it is a microlanguage on its own. To say that it is two different dialects, one a dialect of Ukrainian and the other a dialect of Belarussian, seems irrational in light of the fact that it is obviously a single entity.

Some scholars believe that West Polesian is a separate language (Moseley 2010), but the case has not yet been proven. Another opinion is that this is a Belarussian dialect that extends into both countries.

None of the treatments of Polesian so far seem satisfactory, and it seems to be stuck in classification limbo.

Eastern Polissian or East Polesian is spoken in Ukraine in Chernihiv except for the southeastern districts, in the northern part of Sumy, and in the southeastern portion of Kiev Oblast and in Russia in the southwest part of Bryansk Oblast around Starodub and some parts of Kursk, Voronezh, and Belgorod Oblasts. The lexis looks more Russian as it gets closer to the Russian border. This dialect is best seen as a Ukrainian dialect transitional to Russian. Both Russian and Ukrainian

grammars are used.

Ukrainian intelligibility of other Slavic lects:

There were a number of Ukrainian informants, however, an element of nationalism caused some problems. There were no personal Ukrainian informants.

Ukrainian has full intelligibility of Balachka at 90%. Balachka is best seen as a Ukrainian dialect, while at the same time being part of Eastern Ukrainian.

Ukrainian has [85% intelligibility](#) of the Surzyk mixed language spoken in Ukraine.

Ukraine has [85% intelligibility](#) of Eastern Ukrainian.

Ukrainian has 75% intelligibility of Belarussian (Mezentseva 2014).

Ukrainian also has good intelligibility of Podlachian, possibly 75%.

Ukrainian intelligibility of Lemko Rusyn is 66%. It varies from zero to near-full intelligibility (the latter coming from Ukrainian nationalists, so it is quite suspect). The best solution seems to be to say that Ukrainian is not always fully intelligible with Lemko. This was formerly part of the dialect continuum between Ukrainian and Polish, and Ukrainian has 32% intelligibility of Polish, and halfway would be 66%, so Ukrainian intelligibility of Lemko Rusyn may be 66%.

Lemko is certainly not a Ukrainian dialect as many Ukrainians say, as the only Ukrainians who can understand it are those who were exposed to it via grandparents, and even those can understand but not speak Lemko. There is no such thing as a dialect of X language that X speakers cannot understand, therefore Ukrainian and Lemko Rusyn are separate languages.

Ukrainian understanding of spoken Bulgarian is not known, but [written intelligibility](#) is high, possibly 60%.

Ukrainian has 50% intelligibility of Russian. Ukrainian has excellent [written intelligibility of Russian at 85%](#). Nevertheless, there is a strong factor of bilingual learning here. We can see this when we look at the figures for Canadian Ukrainian, which has only 5% understanding of Russian. This is important because apparently they have had little to no exposure to Russian, while Ukrainians are inundated with the language.

The Russian language in the Ukraine has been declining recently, mostly because since independence the authorities have striven to make the new Ukrainian as far away from Russian as possible.

Hence, Russians understand the colloquial Ukrainian spoken in the countryside pretty well, but they understand the modern standard heard on TV much less. This is because colloquial Ukrainian is closer to the Ukrainian spoken in the Soviet era, which had huge Russian influence.

In recent years, even the Standard Ukrainian has been rapidly changing. Some Russians said they could understand Standard Ukrainian fairly well in the mid-2000's, but since then there has been so much Polonization of the language that Ukrainian is hard for them to understand, and it no longer even sounds like an East Slavic language.

It is often said that Ukrainian and Russian are intelligible with each other or even that they are the same language (a view perpetuated by Russian nationalists). This view is false and is a product of Russian nationalism.

Ukrainians can understand Russian much better than the other way around, but this may be complicated by very heavy bilingual learning on the part of Ukrainians. Ukrainian intelligibility of Russian is hard to measure, as Russian is widely spoken as a second language in Ukraine, and even if they don't speak it, Ukrainians have massive exposure to the language. The situation is similar to the role of Czech in Slovakia.

Most of the Ukrainian speakers who do not speak Russian are in Canada at the moment. Ukrainian-speaking Canadians say they do not understand Russian at all. Therefore inherent Ukrainian-Russian intelligibility may be much lower than as generally thought, and the higher numbers from Ukrainians are due to bilingual learning and heavy exposure to Russian. A good Ukrainian-Russian MI is not available due to massive bilingual learning. The best place to study the inherent MI of these languages is among Ukrainian speakers in Canada.

Ukrainians understanding of Eastern Slovak may be 50%.

Ukrainian intelligibility of Slovak is quite good at 45%. Ukrainian has a lot of Slovak influence compared to Russian, which has almost none.

Ukrainian written intelligibility of Slovak is excellent at 90%.

Ukrainian intelligibility of Rusyn is 39%. It is 75% for the most commonly spoken somewhat watered-down form, but the hard Rusyn spoken by some old men is barely understood at 7%.

There is good evidence that intelligibility between Ukrainian and some forms of Rusyn may be greatly exaggerated.

We have excellent data from the Hutsul dialect of Rusyn on Ukrainian of Rusyn. Hutsul is the Rusyn dialect that may be most closely related to Ukrainian, yet Hutsul-Ukrainian intelligibility can be very poor. The hard, pure Hutsul can be understood by Ukrainians only with great difficulty, and many Ukrainians say they cannot understand it at all. Even the less pure, more watered-down variety spoken in towns can be hard to understand. Hutsuls who lead Hutsul tours say they have to speak in Ukrainian because when they speak in Hutsul, Ukrainians barely understand a word (Coyne 2014).

Ukrainians have 7% intelligibility of pure Hutsul but 70% intelligibility of the Ukrainianized Hutsul (now pretty much Standard Hutsul). The median is 39%.

Ukrainian intelligibility of Polish is 35%. Written intelligibility is higher at 50%.

Ukrainian intelligibility of Kashubian is low at 25%.

Ukrainian intelligibility of Serbo-Croatian is similar to how well Russians understand that language; it may be 20%.

Ukrainian intelligibility of Slovak is poor for some reason at 20%. This is hard to explain. However, written understanding is full at 90%, which seems very hard to explain.

Western or Galician Ukrainian, Standard Central Ukrainian or Ukrainian, and Eastern Ukrainian are the three separate languages that may make up what we think of as Ukrainian. Eastern Ukrainian would also include Kuban Russian and Balachka spoken over the border in Russia.

Such a split would be justified on intelligibility grounds at least.

Western or Galician Ukrainian, encompassing the Galician dialects such as Upper Dniestrian and other Galician dialects, should perhaps be considered a separate language on structural grounds and due to difficult intelligibility with Eastern Ukrainian.

Eastern Ukrainian, encompassing Kuban Russian/Balachka and Slobozhan Ukrainian, should be possibly be thought of as another completely separate language, also on structural grounds and due to problematic intelligibility with Western Ukrainian.

I realize this decision will be very controversial, but I feel it is justified.

Of course then, like Kashubian, we would need a Standard Ukrainian in the middle encompassing everything that is not Rostov/Donbass in the east or Galician in the West, hence the mandate for separating Ukrainian Proper from the far east and west.

The rationale for making such a split will now be discussed.

Western Ukrainian

Western Ukrainian, which may be a separate language, consists of *Upper Dniestrian Ukrainian* and other Western Ukrainian dialects spoken in the far west of Ukraine in the Lviv Ternopil and Ivano-Frankivsk Oblasts in and around Lvov in Galicia. Upper Dniestrian is influenced by German and Polish. This is the main dialect spoken in Galicia.

Nowadays the more pure Western Ukrainian is spoken in the rural areas, as the speech of the cities, even Lviv, have shifted over more towards the standard. In addition, many people in Lviv now speak Russian. Standard Ukrainian speakers refer to this speech as something along the lines of "hick, low class, poor, dumb farmer, etc. speech."

This language has very heavy Polish influence on the lexicon. In fact, outside of Lviv, there are still a number of villages where many people still speak Polish.

Western Ukrainian and Eastern Ukrainian are so far apart that they are nearly separate languages in an objective sense.

In fact, the implicit argument that they are separate languages has already been made by prominent Slavacists (Lunt 1997). At this point, I feel something radical should be done.

Western Ukrainian intelligibility of other Slavic lects:

There were no Western Ukrainian informants at all, making the study of this language most challenging.

Western Ukrainian may have 79% intelligibility of Lemko Rusyn, a Polonized form of Rusyn. Western Ukrainian is a heavily-Polonized form of Ukrainian.

Western Ukrainian has 75% intelligibility of Rusyn. Ukrainian is close to Rusyn, but intelligibility of Rusyn and Ukrainian is much exaggerated. For pure Rusyn, even Western Ukrainian speakers struggle to understand it (Coyne 2014). It works the other way too. In the early part of this century, newspapers printed in Western Ukrainian were distributed in the Rusyn area, however, Rusyns complained that they could not understand the articles.

The marginal ([possibly 60%](#)) intelligibility between Western/Galician Ukrainian and Eastern Ukrainian/Kuban Russian/Balachka/Slobozhan is grounds to split Ukrainian into two languages on the basis at least of MI.

Western Ukrainian may have 57% understanding of Polish, much higher than for Ukrainian. Western Ukrainian is heavily-Polonized, so it makes sense that it would understand Polish much better than Ukrainian Proper.

Western Ukrainian also may have 57% comprehension of Eastern Slovak.

Western Ukrainian may have good intelligibility of Slovak of 40%.

Western Ukrainian may have only fair intelligibility of Russian at 30%.

Eastern Ukrainian

Eastern Ukrainian, possibly a separate language, consists of Slobozhan or Slobodan and Balachka or Kuban Russian.

Slobozhan is spoken in the far east of the country in Kharkiv, Lugansk, and Donetsk. This is a Ukrainian-Russian transitional dialect that is hard to classify.

Slobozhan Ukrainian is the name of this dialect on the Ukrainian side

of the border, where it is best seen as a Ukrainian dialect that is transitional to Russian. 25 years ago, it was much more Russified than it is now as it has become more Ukrainian over time. But Slobozhan Ukrainian gets [more Russian](#) as it moves towards the north and east.

Slobozhan Russian is the name of the dialect after it crosses the Russian border. This dialect is spoken in and around Kantemirovka in the Voronezhskaya Oblast or Voronets Region of Russia near Krasnodar, Belgorod, and Voronezh. The truth is that Slobozhan Ukrainian and Slobozhan Russian are better seen as part of a single language, a Ukrainian-Russian microlanguage called Eastern Ukrainian that is not a part of either Russian or Ukrainian per se but is much closer to Ukrainian than to Russian. There is [no clear line](#) between Ukrainian and Russian in this area, as the Russian language here is quite Ukrainian-like, and the Ukrainian language is quite Russian-like.

Surzyk is the name for both Slobozhan and Balachka, the Central Ukrainian dialects of Cherkasy, Poltava and Zaporozhia, and even Rusyn. It is also the term used for the mixed Russian/Ukrainian speech commonly spoken in the Ukraine. More generally, it is also used for any mixed Ukrainian speech such as Rusyn, West Polissian, etc.

Surzyk intelligibility of other Slavic lects:

Surzyk has 85% intelligibility of both Ukrainian and Russian.

Eastern Ukrainian intelligibility of other Slavic lects:

There were no Eastern Ukrainian informants at all, making MI studies quite trying.

Intelligibility of Eastern Ukrainian and Surzyk in general with Ukrainian is [85%](#).

Eastern Ukrainian intelligibility of Russian is [estimated at 70%](#).

Eastern Ukrainian comprehension of Western Ukrainian may be as low as 60%. Eastern Ukrainian speakers find it [easier to understand their Russian neighbors](#) than Upper Dniestrian and other Galician speakers.

Rusyn

Rusyn is a full language, not a dialect of Ukrainian. The language is composed of 50% Slovak roots and 50% Ukrainian roots, so some difficult intelligibility with Ukrainian might be expected.

It is a mixture of Eastern Slovak dialects with heavy East Slavic influence.

Rusyn is similar to Eastern Slovak, but it has East Slavic phonology and vocabulary (Martin 2013). It has also been described as either Polish-Slovak transitional, Ukrainian-Slovak transitional or possibly both.

The standard Ukrainian view is that Rusyn is a Ukrainian dialect, and the standard Slovak view is that it is a Slovak dialect (Kushko 2007). Neither are correct, but some lects classed as Rusyn are actually part of Eastern Slovak.

Rusyn began breaking away from the rest of East Slavic 600 years ago, since 1400 (Kushko 2007). That surely seems to be enough time to split into a separate language.

There are four main Rusyn dialects: Transcarpathian Rusyn, spoken in Ukraine, Poland, Hungary and Romania; Lemko Rusyn, spoken in Poland; Boyko; and Hutsul, the latter two spoken in the Ukraine. (Kushko 2007).

Vojvodinian, Pannonian, or Barachka-Srem Rusyn, spoken in Yugoslavia is not even a part of the Rusyn language; instead it is a separate language that is part of Marco-Eastern Slovak (Lund 1997).

Slovakian Rusyn also appears to be a part of Eastern Slovak (Vanko 1997). Even linguists who say that Rusyn is a Ukrainian dialect often agree that Pannonian Rusyn and Slovakian Rusyn may not be a part of Rusyn.

Lemko has Ukrainian, Slovak, and Polish influences; Slovakian Rusyn has influences mostly from Slovak but also partly from Czech; Transcarpathian Rusyn has Ukrainian, Slovak and Hungarian influences; and Pannonian Rusyn has influences from Hungarian and Serbo-Croatian (Kushko 2007).

Intelligibility within Rusyn is generally described as full, and they are said to communicate together easily.

However, other Rusyns say that every Rusyn speaks a different language depending on where they came from.

Others say that if you want to learn Rusyn, any given book that you find will be about only one of the dialects. But if you learn one dialect, you will have an easier time learning the others.

Both reports imply that Rusyn internal intelligibility may be somewhat less full than most people think.

The divisions in the Rusyn language are significant enough to have impeded the development of a Rusyn literary language, as literature produced in one dialect would be rejected by the others who said it was not their speech. For instance, there are three quite different words for the word *verb* - one in Subcarpathian and Lemko, a second in Slovakian Rusyn, and a third in Pannonian Rusyn (Maksimovich 2008).

Hence, instead of one standard Rusyn languages, there are four standards – Standard Lemko, Standard Slovakian Rusyn, Standard Subcarpathian Rusyn, and Standard Pannonian Rusyn (Vanko 1997). However, Slovakian Rusyn and Pannonian Rusyn are part of Eastern Slovak, so we are comparing two different languages here, and therefore, this contrast shows us nothing about intelligibility within Rusyn Proper.

Structure is one way to separate a language from a dialect. There are opinions suggesting that Lemko and Boyko are structurally separate languages and hence are full languages on that basis. However, this argument is not yet proven.

Transcarpathian, Subcarpathian, Dolyniane, Dolyshniany, or Lowlander Rusyn spoken in Ukraine, Poland, Hungary and Romania, which includes Standard Rusyn, is usually seen as a dialect of the Rusyn language. Some say that this lect is fully intelligible with Ukrainian, but a closer examination shows that this is not the case. Intelligibility instead is probably more on the order of Ukrainian-Belarussian, or 75-80%. In some cases, it may be much lower than that. Speakers are called Dolyniane, Dolyshniany, or Lowlanders and represent the largest of the Rusyn groups.

There are several subdialects in this group such as *Northern Maramoroš, Southern Maramoroš, Uz, and Bereg* dialects (Vanko 1997).

Most live in Ukraine, but a few live in Slovakia and Romania, with a tiny number living in Hungary.

A Transcarpathian Rusyn standard language for the Rusyns of the Ukraine was codified in 1999 based on the Southern Maramoroš dialects balanced with elements of the Northern Maramoroš, Southern Maramoroš, Uz, and Bereg dialects and the East Zemplyn dialect of Slovakian Rusyn (Vanko 1997).

Russians have high intelligibility of Rusyn, much higher than for Ukrainian. Ukrainian is much harder for Russians to understand than Rusyn. Russians and Ukrainians understand it so well because this group of Rusyns still speaks the old language of the Kievan Rus, which this area was a part of.

When the Rus broke up, its residents dispersed in many directions, and those that remained continued speaking the Kievan Rus dialect which evolved into the Rusyn language. This implies that the old language of the Kievan Rus is still highly intelligible to Russian and Ukrainian speakers.

The Kievan Rus-derived Rusyn is dramatically less Polonized than Standard Ukrainian, and this is why Russians can understand it well. The fact that this dialect is understood much better than Ukrainian by Russians leads Russians to call Rusyn a form of *Surzyk*. *Surzyk* is a Ukrainian-Russian mixed language spoken in Ukraine. Rusyn seems similar to *Surzyk* because to Russians it sounds like Ukrainian with a heavy mix of Russian in it. The Russian element then is from the language of the Kievan Rus.

Hungarian Rusyn is the Transcarpathian Rusyn spoken by the Rusyns of Hungary living in a few villages in Northeastern Hungary (Vanko 1997). Their dialect is quite a bit different from that of the Ukrainian Transcarpathian Rusyns. It has a lot of Hungarian influence in it.

Gergely Benedek, a Hungarian researcher of the Hungarian Rusyns, created a codified Hungarian Transcarpathian Rusyn language based on the *Komloška* dialect in 2004 (Vanko 1997).

Romanian Rusyn is the Transcarpathian dialect spoken in a dozen villages of Rusyn speakers in the Maramoroš region of Northcentral Romania along the Ukrainian border (Vanko 1997).

Lemko Rusyn speakers are now split about equally between Slovakia and Poland.

53% of the Lemkos live in Poland, 25% of them in their former homeland in Lemkivschyna in the Carpathians and another 75% in

northern Poland where they were resettled during Operation Vistula, a violent counterinsurgency against the Ukrainian Insurgent Army during World War 2 (Maksimovich 2001). The total population of Lemkos has been reduced from ~195-200,000 before World War 2 (70% in Poland) to its present population, which is ~120,000. This represents a decline in population of 39% in the last 70 years.

The causes of this population loss and what happened to the missing people are not known.

The best guess is that most of them assimilated to Polishness and dropped their Lemko identity (Maksimovich 2001).

47% of the Lemkos reside in the Preshov Region of Slovakia adjacent to Ukrainian Carpathia (Kushko 2007). Only 4% of Lemkos remain in their former homeland in Polish Carpathia.

Lemko is different from Transcarpathian Rusyn, Hutsul, and Boyko (Kushko 2007), but the extent of this difference, and MI between Lemko and the three other Rusyns is not known. It differs from Transcarpathian Rusyn in that it has a lot of Polish vocabulary and Slovak influences, whereas Standard Rusyn has more influences from Hungarian and Romanian.

Although this is traditionally seen as part of the dialect continuum between Ukrainian and Polish, it is better seen as a Polish-Slovak transitional lect. In fact, the Slovak influence is so strong that some even consider it to be an Eastern Slovak dialect like Slovakian Rusyn and Pannonian Rusyn. At the moment, there is no good evidence to lump Lemko into Eastern Slovak with the other two Rusyn dialects.

Some sources feel that Lemko is a full language in its own right, but there is no good evidence for this yet.

Lemko has official status in Poland.

A standard Lemko language was codified in 2000 (Chomiak and Fontanski 2000).

Lemko intelligibility of other Slavic lects:

There were no Lemko informants at all, making the study of Lemko MI nearly impossible.

Lemko may have 60% intelligibility of Ukrainian.

Hutsul Rusyn is spoken mostly in the Ukrainian Carpathians and the

northern extremes of Romania in Bukovina and Maramureş (Kushko 2007). 83% of Hutsuls live in Ukraine, and only 17% live in Romania. Hutsul is very similar to Ukrainian with some Polish influences.

Hutsul also has a lot of Romanian loans (Coyne 2014).

Most young people can speak Hutsul. But instead of speaking pure Hutsul, they speak a sort of "Hutsul Surzyk" composed of half Hutsul and half Ukrainian. This dialect encompasses their identity as well, as the young people see themselves as both Ukrainian and Hutsul. This is now the preferred language of the young people in the region (Coyne 2014).

Hutsul itself is undergoing a lot of contact-induced changes from Ukrainian.

These changes were ongoing even as far back as 50 years ago (Rudnyckyj 1965). Present Hutsul speech now has many Russianisms in it in addition to many Ukrainian-induced changes in phonology (Lesjuk 2012).

Hutsul is in danger of extinction (Coyne 2014). All education has been in Russian since the USSR annexed the region in 1945, and Ukraine considers Hutsul to be a Ukrainian dialect, which is not true.

The purest Hutsul is spoken high up in the mountains by old people who are about 80 years old. Some Hutsuls give the dialect only 20-30 more years until it dies (Coyne 2014).

However, as children are still being raised speaking Hutsul, the dialect should last another 80 years at least in its impure Surzyk form.

In fact, in the Verxovyna area, most children still show up to school as Hutsul monolinguals (Coyne 2014).

The young people do not see Hutsul as a language with much value. They see Ukrainian as the ticket to occupational and financial reward (Coyne 2014). Young people have been abandoning Hutsul for Ukrainian for 50 years now. Ukrainian was seen as the language of school, theater, books, and journals (Rudnyckyj 1965).

In addition, there are now many mixed marriages between Hutsuls and Ukrainians. In these marriages, Ukrainian or even Russian is the language of the household (Coyne 2014).

There are now movements to revive the Hutsul dialect, but how well these will be implemented, whether they will be implemented at all,

or what success they will have if they are is very much up in the air. These movements are having problems because Ukraine continues to insist that Hutsul is a Ukrainian dialect (Coyne 2014).

Boyko Rusyn is spoken to the north in the Ukrainian Carpathians and across the border in Poland in Subcarpathian Voivodship. The Boykos were affected by the same counterinsurgency that devastated the Lemkos, and many were deported out of Poland to the USSR after World War 2, after which their property was stolen by Poles. Boykos regard the term *Boyko* as pejorative. On censuses, they identify as Ukrainians. The Boyko dialect is heavily based on OCS.

A few Soviet scholars regarded the Boykos as a specific ethnic group speaking their own language, of which *bo ye* "yes" is an identifying feature (Nikitin et al 2009). At present, there is no good reason to separate the Boyko dialect from the rest of Rusyn.

But there is good reason to see the Boykos as genetically separate from the Lemkos and Hutsuls (Nikitin et al 2009).

Rusyn intelligibility of other Slavic lects:

There was a serious dearth of Rusyn informants, hence a lot of the judgments were simply guesswork.

Rusyn-Ukrainian intelligibility is 78%. The MI between these two languages is described as similar to Ukrainian-Belarusian intelligibility (78%), so Ukrainian and Rusyn may have very good but not full intelligibility at 75-80%.

However, the Hutsul data above suggests that Ukrainian intelligibility with Rusyn is greatly exaggerated, as Ukrainians cannot even understand the most Ukrainian-like Rusyn dialect, Hutsul, well or even at all (Coyne 2014).

In the early 1900's, there were many "Ukrainianizers" in the Rusyn area in the Subcarpathians who were trying to get the Rusyns to speak and write in Ukrainian. These Ukrainians insisted that Rusyn was simply a Ukrainian dialect and that Rusyns simply needed to learn to read, write, and speak the proper language. Rusyn newspapers at the time decried this trend, printing articles opposing it. The articles opposed the efforts of the Ukrainianizers, whom they said were promoting a foreign language (Rusinko 2003).

They also disliked the Ukrainian articles increasingly being seen in the

Rusyn press. The articles were written in the Galician dialects of Western Ukrainian, which the Rusyn press called a "half-Polish language." The articles also noted that many Rusyns were complaining that they could not understand articles written in Galician Western Ukrainian. Based on this evidence, it appears that Rusyns cannot understand even written Western Ukrainian. They must understand the spoken language even worse (Rusinko 2003).

Rusyns have good intelligibility of Russian at 75%. The reasons are described in the main Rusyn entry above.

Rusyns also have good understanding of Polish at 75%.

Rusyn also has [some intelligibility](#) of Eastern Slovak. 50% is a good estimate. Eastern Slovak-Rusyn MI is limited in that technical terminology and upper register expressions are different.

Russian

Russian is a very old language. The earliest records of what looks like Russian are from 1100. There are written documents from Russia dating as far back as 800 written in Slavic, but this is more Proto-East Slavic than anything else. The base of Russian is the liturgical language Old Church Slavonic (OCS), which was a koine widely understood by Slavic speakers at the time when Proto-Slavic was breaking up into East, West and South Slavic around 500. In fact, Russian's Cyrillic script was created for the specific purpose of writing OCS.

In the 1200's, the Rus was conquered by Mongols which brought many Turkic loans into the languages, mostly dealing with matters of state and military.

Up until the 1300's, the ancestors of the Ukrainians, Belarussians and Russians in the Kievan Rus all spoke dialects of Old East Slavic, which existed alongside the literary language of OCS. However, birch bark writings from Russia in the 1300's show that people were already starting to write in their Old East Slavic vernaculars by this time.

In the 1300's, the Mongol Empire broke up, and the area was split into two states, a Grand Duchy of Moscow to the East and a Grand Duchy of Lithuania to the West. The dialects of the west came to be

called Ruthenian and came under heavy Polish influence. In the East, the language continued to be influenced by OCS. Ruthenian would later develop into Ukrainian, Belarussian, and Rusyn. The true base of Ruthenian became Rusyn, while Ukrainian and Belarussian branched off.

In Russia, there developed a war between the vernacular and the literary language, OCS. With the fall of the Byzantine Empire in the 1450's, Bulgarian and Byzantine scholars moved to Moscow, where they were upset to find that OCS was becoming Russified by the vernacular.

They initiated a standardization effort to remove what they saw as the impurities of the vernacular. This had the result of pulling the language back towards the literary standard.

Under Peter the Great around 1700, the written language was simplified to increase literacy among Russians by eliminating some letters and diacritics. He also tried to modernize and Westernize the country and orient it more towards Europe. French then became the spoken language of the cultured upper classes. Peter also authorized the importation of many European words, especially from French, Dutch, and German.

As a result, Russian was being flooded with French, Latin and German words. This resulted in a Russification project to rid Russian of West European borrowings and replace the borrowed words with Slavic words. In order to coin new words, Russian scholars searched OCS and replaced many of the Western borrowings with OCS words.

Nevertheless, Russian still has 10% each French and German borrowings, more than in most Slavic languages other than East Slovak, Silesian, and Kaikavian. The other 60% of Russian is Old Church Slavonic. Only 10% of Russian is directly cognate with Ukrainian and Belorussian which explains Russian's lower MI with Ukrainian.

In 1775, Mikhail Limonosov wrote a grammar of Russia which for the first time married OCS, which he called high style, with the vernacular, which he called low style. He suggested that a middle style, a blend of the two, would be best for written Russian. A number of writers then began to use this middle style to create what became Modern Standard Russian.

Alexander Pushkin was particularly important in this regard as he used all three styles in his writing, which also made the vernacular a legitimate form of Russian writing style. He also directly translated many phrases and figures of speech from European languages, which greatly expanded Russian vocabulary.

In 1882, a Danish linguist named Dahl wrote a dictionary of Russian, devoting 50 pages to Russian dialects. Each province had a different dialect, with the dialects far enough apart to be considered separate languages. Those 50 pages were removed from the dictionary by Czarist censorship. Over the next 150 years, almost all of these diverse dialects would be assimilated to a Standard Russian language.

The Novgorod dialects in the area of what is now Belorussia and adjacent Russia had many Ukrainian words. But this dialect has been dead for a long time ago, though it was once the spoken language of Novgorod. These words were left by Ukrainian traders moving through the region. Dahl also noticed similarities between Novgorod Russian and the Southern Russian Cossack dialects from Peskov to Don, also heavily Ukrainized.

What is interesting about the Novgorod dialect is that it may be the remains of what was once North Slavic, a branch of Slavic never thought to have existed. Anatoli Zhuravlyov has found that Novgorod Russian retained several Proto-Slavic forms that have been lost in all other branches. This implies that Novgorod Russian was the remains of a fourth branch of Slavic now gone extinct – North Slavic. This theory was the subject of a Netflix TV Countrywide TV show.

Published linguistic documents make reference to a North Slavic branch, citing specific sound changes that occurred in this group (Kortlandt 2003), which implies that its existence is more of an established fact as opposed to a theory.

These documents place the North Slavic changes all the way back to Early Proto-Slavic (Kortlandt 2003), which means North Slavic may have been one of the earliest branches of Slavic, dating even before Early Middle Slavic. This period is dated from 300-500 (Kortlandt 2003).

Russian shares a great deal of common vocabulary with the rest of Slavic (Sexton and Shapoval 2002).

Russian intelligibility of other Slavic lects:

The study of Russian MI was helped by the presence of many Russian informants. There were no usable personal informants, as Russian linguists who were interviewed declined to be quoted. Nevertheless, their judgments were used anonymously.

Russian [intelligibility of Surzyk is 85%](#) .

Russians have 85% intelligibility of Ninilchik Russian (Leman 2016).

Russian intelligibility of Belarussian is no higher than 75%, and in some cases even lower, depending on the Russian dialect the Russian speaker is speaking (Mezentseva 2014). Russian and Belarussian share a very high level of vocabulary at 92%. Russians understand Belarussian [very well when it is written, 85%](#).

Russian [intelligibility of East Ukrainian is 70%](#).

Russians understand 65% of Bulgarian but the range is very high and some Russians understand Bulgarian poorly. However, Russian understanding of written Bulgarian is very high at 90%.

Russians understand only 60% of Balachka, even though it is spoken in Russia.

Russian intelligibility of Novgorod Russian is [less than 50%](#).

Russian has only 40% inherent intelligibility of Ukrainian. However, Russians have [excellent understanding of written Ukrainian at 90%](#).

Ukrainian-Russian inherent MI may be extremely low, near zero. Ukrainian-Russian inherent intelligibility may be 5% based on figures from Canadian Ukrainians who are not exposed to Russian. Similarly, Russian inherent intelligibility of Ukrainian may be very low. Russians who have not been exposed to Ukrainian in a long time [may only understand 5%](#) of Ukrainian spoken at normal speed. This is the same figure for Ukrainian-Russian intelligibility, so Ukrainian-Russian MI may be only 5%.

So it is not true at all as it often said that Ukrainian and Russian are mutually intelligible. For example, all Russian shows get subtitles on Ukrainian TV. Yet some say that the subtitles are simply put on as a political move due to Ukraine's puristic language policy. The situation in Ukraine does not reflect Ukrainian-Russian inherent MI well due to massive bilingual learning on the part of Ukrainians.

In addition, Russians say that they used to understand Ukrainian well 10 years ago, but in the last decade, the language has changed so

much that Russians now say it is not even Ukrainian anymore, and they say Ukrainians speak Polish. Many Russians say that they cannot understand the new Polonized Ukrainian well.

Russians only can understand a meager 50% of Volgograd Russian, which is an actual dialect of the Russian language spoken 300 miles northeast of Moscow.

Russian intelligibility of Rusyn may be 45%. Rusyns say that Russians, at least those around Ryazan' or Nizhniy Novgorod 200-300 miles south-southeast of Moscow, often do not understand them at all. Other Russians have 90% intelligibility of Rusyn, so the figure ranges from 0-90%, with 45% being the medium.

Russians understand Eastern Slovak well at 45%.

Russians have good intelligibility of Čakavian at 45%, which seems strange, but Čakavian has excellent intelligibility of Old Church Slavonic (OCS) on which Russian is based.

Russians understand Slovak well at 42%.

Russians understand the Banat Bulgarian spoken in Romania at 40%, less well than they understand Bulgarian. The sample size here was small.

Russian comprehension of Croatian is 35%, about the same as with Serbo-Croatian.

Russians understand Pannonian Rusyn, actually an Eastern Slovak dialect, at 32%, but the sample size was very small.

Russians also have 32% of Western Ukrainian, a bit better than for Polish.

Russian intelligibility of Serbo-Croatian is 30% for oral and 50% for written Serbo-Croatian (Mezentseva 2014).

Russian only has fair intelligibility of Slovak, maybe 30%, but written is higher at 63%.

Russians have 27% of Macedonian, a lot lower than their understanding of Bulgarian. The reason for the difference is not clear.

Russian has 18% comprehension of Serbian, quite a bit lower than for Croatian. This is odd as Serbian has quite a bit of OCS in it, and there are many words that Serbian shares with Russian that are not in Croatian.

Russian may have 25% intelligibility of Polish. Written intelligibility of Polish is higher at 70%. Russians have a much harder time understanding Polish than Ukrainians and Belarussians do. The grammar is quite similar, but the phonology and the lexis are very different. In terms of lexis, Polish has many words of Latin origin that Russian lacks. But Russians who understand Ukrainian can understand much more Polish, possibly 40%, as Ukrainian could be seen as a Polish-Russian mixed language (Mezentseva 2014).

Russians understand very little of Upper Sorbian, 17%.

Russian understanding of Slovene is very poor at 10%. Written is quite a bit better at 25%.

Russians comprehend Kashubian very poorly at 8%.

Russians understand almost nothing of Czech, 4%. Oddly enough, Russians understand written Czech very well, 70%.

Volgograd Russian is a highly divergent Russian dialect spoken in Volgograd 300 miles northeast of Moscow. Russian intelligibility of this dialect is quite low at less than 50%. This may be the remains of a dialect spoken traders along the Northern Trade Route that extended from the Bay of Finland to the Taz River in Russia. The trade went along Russia's northern coast via the White Sea, the Barents Sea, and the Kara Sea. The trade went along Russia's northern coast to Northern Europe and back. Many European loans went into this dialect. The trade route was shut down in the 1660's due to fears of European infiltration into Russia.

Further study is needed to see how prevalent this difficult intelligibility is. Splitting off as a separate language would cause a firestorm, so for the moment, Volgograd Russian is best seen as a divergent Russian dialect.

Ninilchik Russian

Ninilchik Russian is a separate Russian language, newly discovered, that is spoken by only a few very old people in the small village of Ninilchik on the central coast of Alaska. It is full of loans from Alaskan native languages, and Russian intelligibility of it is not full. Ninilchik

was settled by Russians in 1840 and this is when the language began. For the next forty years, it received no new Russian settlers, so the language became frozen in time. It combines an 1840's-era Russian with a multitude of native loans.

When Alaska became part of the US in 1867, the village hardly blinked. There were many other Russian-speaking settlements in Alaska at that time, but US troops were dispatched to the new state to secure it in the name of the US. At this time, all Russian settlements but Ninilchik and a few others switched over to English. Forty years later, in 1907, Ninilchik villagers were still keeping records with the Julian calendar and speaking only Ninilchik.

But a new settlement school, the Territorial English School, opened up in 1911 and changed everything. All students spoke only Russian, but speaking in Russian was forbidden at school in an attempt to force them to learn English, a common practice at Indian and Aborigine schools in the US, Canada and Australia at the time. Punishment was having your mouth washed out with soap containing naphtha, a toxin.

Many associated Ninilchik speakers with speakers of native languages against whom there was extreme discrimination. In order not to be confused with them, Ninilchik speakers spoke English when they went to big cities. When relations with the USSR warmed in the 1970's, many Soviet citizens visited Alaska. Some came to Ninilchik, where they laughed at the natives' antique language with no words for many modern items and archaic terms for other things now replaced by newer words in Modern Russian.

Finally in 1997, word got around to Russian linguists about this interesting Russian language in Alaska, and two prominent linguists traveled to Ninilchik to begin recording it. Oddly, the language lacks words for common things. Mira Bergelson, one of the Russian linguists, never could find a Ninilchik word for grandchild.

Ninilchik was not the only Russian-speaking village in the area. But the 1964 Earthquake destroyed several nearby Russian-speaking villages. These people also spoke Ninilchik dialects. This dramatically reduced the native speaker population.

Standard Russian speakers from Russia have high but not full intelligibility of Ninilchik. Intelligibility is hampered by the different pronunciation and the occasional native loans. On the other side,

Ninilchik speakers do not fully understand Modern Russian because the pronunciation is so different from what they speak, and there are many modern Russian words that they don't know.

Bergelson, her linguist husband Andrej Kibrik, and cultural anthropologist Nina Raskladkina published the first dictionary of Ninilchik. The dictionary is an ongoing project as it is continually edited and updated. In 1997, the linguists began with 20 speakers. Now they are down to only a few, all over 85. The language is moribund and is expected to go extinct shortly. This is absolutely a full separate language from Russian and it deserves an SIL code, the granting of which should be at least scientifically uncontroversial.

Ninilchik Russian intelligibility of other lects:

Ninilchik MI study was helped by the presence of an excellent linguist who has also extensively studied the language.

Ninilchik Russian has 65% intelligibility of Russian (Leman 2016).

So far, this paper contains 103 references.

References

Albin, Alexander and Alexander, Ronelle. 1972. [The Speech of Yugoslav Immigrants in San Pedro, California](#). Publication of the Research Group for European Migration Problems XVII. Springer Science & Business Media. The Hague: Martinus Nijhoff.

Author Collective. 1885-1892. Meyer's Conversation Lexicon. Leipzig/Vienna: Bibliographic Institute.

Bidwell, Charles. Spring 1966. [The Language of the Backa Ruthenians in Yugoslavia](#). *The Slavic and East European Review* X, p. 32-45.

[Breu, Walter](#). October 1-3, 2009. "Tense, Aspect, and Mood in Molise Slavic," paper presented at the Morphologies in Contact Conference, Universität Bremen, Bremen, Germany.

Breu, Walter. May 2015. Professor of Slavic Linguistics, Linguistics Department, University of Constance, Constance, Germany.

Personal communication.

Budovičová, V. 1987. Semikomunikácia ako Lingvistický Problém (Semicommunication as a Linguistic Problem). J. Mistrík, ed. *Studia Academica Slovaca* 16. Prednášky XXIII (in Slovak).

Bujan, Željka. September 2015. Medjimurje Kaikavian speaker and Professor of Croatian Language. Personal communication.

Casad, Eugene H. 1974. Dialect Intelligibility Testing. Summer Institute of Linguistics Publications in Linguistics and Related Fields 38. Norman, OK: Summer Institute of Linguistics of the University of Oklahoma.

Cerna, Iva and Machalek, Jolana. 2007. *Beginner's Czech*. Hippocrene Books.

Čagalj, Ivana, Aleksandra Ščukanec, and Anita Skelin Horvat. 2014. (Re)constructing Language and Identity: Burgenland Croats in Slovakia. Zagreb: Faculty of Humanities and Social Sciences, University of Zagreb.

Celinič, Anita. 2010. Speech of Posavina Bregi near Ivanič-Grada. *KAJ - Journal for Literature, Art, Culture* 43: 214.

Chomiak, Mirosława and Fontanski, Hendryk. 2000. *Grammar of the Lemko Language*. Katowice: Slask.

Comrie, Bernard and Corbett, Greville. 1998. *The Slavonic Languages*. New York: Routledge.

Coyne, Erin Victoria: Coming Down from the Mountain: Dialect Contact and Convergence in Contemporary Hutsulshchyna. Spring 2014. PhD dissertation. Berkeley: University of California.

Delsing, Lars-Olof. 2007. "Scandinavian Intercomprehension Today," in Ten Thije, Jan D. and Zeevaert, Ludger, eds. *Receptive Multilingualism: Linguistic Analyses, Language Policies, and Didactic Concepts*, pp. 231-246. Amsterdam/Philadelphia: Benjamins.

Dressler, Wolfgang and de Cillia, Rudolf. 2006. "Spracherhaltung, Sprachverfall, Sprachtod (Language Maintenance, Language Decay, Language Death)" in U. Ammon, N. Dittmar, K. J. Mattheier, and P. Trudgill (eds.), *Soziolinguistik* Vol. 3, 2258-2271. Berlin: Walter de Gruyter (in German).

Filipovic, Rudolf. 1996. "Sociolinguistic Conditions for Maintenance

of a Dialect: The Case of the Konavle Dialect in Watsonville, California," in Language Contact across the North Atlantic: Proceedings of the Working Groups Held at the University College, Galway, Ireland, 1992 and the University of Göteborg, Sweden, 1993, Ureland, P. Sture and Iain Clarkson, editors. Tübingen: Niemeyer.

Filipovic, Rudolf. 1997. ["The Struggle to Maintain Croatian Dialects in the United States,"](#) in Language and Its Ecology: Essays in Memory of Einar Haugen, Einar Ingvald Haugen, ed. Berlin: Walter de Gruyter.

Fishman, Joshua and Garcia, Ofelia. 2011. Handbook of Language and Ethnic Identity: The Success-Failure Continuum in Language and Ethnic Identity Efforts, p. 367. Oxford: Oxford University Press.

Fraenkel, Gerd. 1962. Mutual Intelligibility between Turkish of Turkey and Azerbaijani, in American Studies in Altaic Linguistics, ed. Nicholas Poppe. Bloomington: Indiana University Press.

Gerdzhikov, Georgi. 2003. Преизказаването на глаголното действие в българския език (Re-declaration of Verbal Actions in Bulgarian). Sofia, Bulgaria: St. Kliment Ohridski (in Bulgarian).

Glasgow, Holly Hood. 2012. Croatian Language and Cultural Maintenance in the Slavic-American Community of Strawberry Hill. MA Thesis. Kansas City, Kansas: Anthropology Department, University of Kansas.

Golec, Boris. 2012. Nedokončana Kroatizacija delov Vzhodne Slovenije Med 16. in 19. Stoletjem. Po Sledih Hrvaškega Lingvonima in Etnonima v Beli Krajini, Kostelu, Prekmurju in Prlekiji (Incomplete Croatization of Parts of Eastern Slovenia during the 16th and 19th Century. Following the footsteps of the Croatian Lingonyms and Ethnonyms in Bela Krajina, Kostel, Prekmurje and Prlekija.). Ljubljana, Slovenia: Založba ZRC (Research Centre of the Slovenian Academy of Sciences and Arts). (In Slovenian.)

Golubovic, Jelena. November 2013-May 2015). Serbo-Croatian speaker, PhD student in Linguistics, University of Groningen. Personal communication.

Golubovic, Jelena and Gooskens, Charlotte. February 15, 2014 (Golubovic and Gooskens 2014a). "How Well Can Speakers of Serbian Understand Other Slavic Languages?," paper presented at

the 13th Annual Conference of the Slavic Cognitive Linguistics Association, Cambridge, Massachusetts.

Golubovic, Jelena and Gooskens, Charlotte. September 11, 2014 (Golubovic and Gooskens 2014b). Mutual Intelligibility in the Slavic Language Area Measured by Cloze Test. *Methods in Dialectology XV*. Groningen, Netherlands.

Golubovic, Jelena and Gooskens, Charlotte. September 11, 2014. (Golubovic and Gooskens 2014c). "Mutual Intelligibility in the Slavic Language Area," paper presented at the Annual Meeting of the Societas Linguistica Europaea (SLE), Poznan, Poland.

Golubovic, Jelena and Gooskens, Charlotte. 2015a. (Golubovic and Gooskens 2015a). Mutual Intelligibility between West and South Slavic Languages. *Russian Linguistics* 39:2.

Golubovic, Jelena and Gooskens, Charlotte. May 28, 2015. (Golubovic and Gooskens 2015b) "Mutual Intelligibility between Czech and Croatian – A Didactic Approach," paper presented at the Receptive Multilingualism: Multilingual Resources in Service of Mutual Understanding Conference, Joensuu, Finland.

Gooskens, Charlotte. 2007. The Contribution of Linguistic Factors to the Intelligibility of Closely Related Languages. *Journal of Multilingual and Multicultural Development* 28: 445-467.

Gradišćansko-Hrvatski Centar (Burgenland Croatian Center). January 3, 2011. The Language of the Burgenland Croats.

Greenberg, Marc L.. 2000. A Historical Phonology of the Slovene Language. Heidelberg: Winter.

Hannan, Kevin. 1996a. Borders of Language and Identity in Teschen Silesia. Berkeley Insights in Linguistics and Semiotics. New York: Peter Lang.

Hannan, Kevin. Winter 1996b. The Lachian Literary Language of Óndra Łysohorsky. *The Slavic and East European Journal* 40:4, 726–743.

Haugen, Einar. 1966. Semicommunication: The Language Gap in Scandinavia. *Sociological Inquiry* 36, 280–297.

Hopkins, Paul. 1996. Kashubian - Language or Dialect? *Working Papers of the Linguistics Circle of the University of Victoria* 13: 51-54

Houtzagers, Peter. 1999. The Kajkavian Dialect of Hidegség and Fertőhomok. *Studies in Slavic and General Linguistics*, Vol. 27. Amsterdam: Editions Rodopi B.V.

Houtzagers, Peter. 2008a. "On Burgenland Croatian Isoglosses," in *Dutch Contributions to the Fourteenth International Congress of Slavists - Ohrid, September 10-16, 2008: Studies in Slavic and General Linguistics*, Volume 34, Janneke Kalsbeek, Peter Houtzagers, and Joe Schaeken, eds. Amsterdam: Editions Rodopi B. V.

Houtzagers, Peter. 2008b. Evidence and Counter-Evidence, *Festschrift Frederik Kortlandt*, Volume 1, SSSL 32, pp. 247-264. Amsterdam: Editions Rodopi B.V.

Houtzagers, Peter. 2014. Burgenland Croats and Burgenland Croatian: Some Unanswered Questions. *Rasprave Instituta 39*: 253-269.

Ivšić, Stjepan. 1911. Prilog za Slavenski Akcenat (A Contribution to Slavic Accent). *Rad JAZU* 187:133–208 (in Croatian).

Ivšić, Stjepan. 1936. Jezik Hrvata Kajkavaca (The Language of Kaikavian Croats). *Ljetopis Jazu*, republished in 1996 by Zaprešić (in Serbo-Croatian).

Jakubinskij, Lav. 1925. Die Vertretung des Urslav ě im Čakavischen (The Reflex of Proto-Slavic ě in Čakavian). *Zeitschrift für Slavische Philologie* 1: pp. 381–396.

Jembrih, Alojz (A. Jembrih 2003) 2003. Obilježja Dvaju Govora uz Rijeku Sutlu: Podčertka u Sloveniji i Razvora u Hrvatskoj (Characteristics of Two Dialects Spoken along the Sutla River: Podčertka in Slovenia and Razvora in Croatia). Zagorje, Croatia: Časopis za Kulturu (in Serbo-Croatian)

Jembrih, Mario (M. Jembrih 2014-2019). Medjimurje Kaikavian speaker, Kaikavian language activist. Msc in Social and Economic Sciences (with heavy emphasis on Linguistics), Vienna University of Technology, Institute for Design and Assessment of Technology, Vienna, Austria. Personal communication.

Jensen, John B. 1989. On the Mutual Intelligibility of Spanish and Portuguese. *Hispania* 72, 848-852.

Jesenský, Marcel. 2014. The Slovak–Polish Border, 1918-1947. New

York: Palgrave Macmillan.

Kalsbeek, Janneke. 1998. [The Čakavian Dialect of Orbaníci near Žminj in Istria](#). *Studies in Slavic and General Linguistics*, Volume 25. Amsterdam: Editions Rodopi R.V.

Kalsbeek, Janneke. 2008. "Past Tense Habitual Expressions in Old Croatian Texts," in *Dutch Contributions to the Fourteenth International Congress of Slavists - Ohrid, September 10-16, 2008*. *Studies in Slavic and General Linguistics*, Volume 34, Janneke Kalsbeek, Peter Houtzagers, and Joe Schaecken, eds. Amsterdam: Editions Rodopi R.V.

Kaplan Robert B. and Baldauf, Richard B. 2005. [Language Planning and Policy in Europe, Volume 2: The Czech Republic, the European Union and Northern Ireland](#). Clevedon, UK: Multilingual Matters.

Kapović, Mate. 2010. [Čiji Je Jezik? \(Who Does Language Belong To?\)](#). (In Serbo-Croatian).

Kapović, Mate. 2017. The Position of Kajkavian in the South Slavic Dialect Continuum in Light of Old Accentual Isoglosses. *Zeitschrift für Slawistik* 2017; 62(4): 606–620. Berlin: Walter de Gruyter.

Karamian, Mark. 2014. Professor, [Dictionary Department, Institute for Linguistic Studies, Russian Academy of Sciences](#), St. Petersburg, Russia. Personal communication.

Kmiecik, Kuśa. November 2014. Polish speaker, graduate student in English and German Philology, Lodz, Poland. Personal communication.

Kolanović, Nikola. August 2015. Southern Čakavian speaker, Croatia. Personal communication.

Kortlandt, Frederick. 2003. ["Early Dialectal Diversity in South Slavic II,"](#) in *Dutch Contributions to the Thirteenth International Congress of Slavists*, Ljubljana: Linguistics (SSGL 30), pp. 215-235. Amsterdam/New York: Rodopi.

Kostomarov, Nikolay. May 2015. ["Knyaz Kostantin Konstantinovich Ostrozhsky,"](#) in *Russian History in Biographies of its Main Figures*. (in Russian).

Kushko, Nadiya. Spring 2007. [Language: The Historical Context and Contemporary Situation](#). *The Slavic and East European Journal* 51:1, pp. 111-132.

- Kustić, N. 2003. Čakavski Govor Grada Paga s Rječnikom (The Čakavian Speech of the City of Pag with Dictionary). Zagreb: Društvo Pažana (Society of Pažana).
- Kutsarov, Georgi. 2007. Теоретична граматика на българския език (Theoretical Grammar of the Bulgarian Language). Морфология. Plovdiv, Bulgaria: Paisiy Hilendarski (in Bulgarian).
- Leman, Wayne. April 2016. Linguist, Ninilchik, Alaska. Personal communication.
- Lesjuk, Mykola. 2012. Dinamika Gucul'skogo Dialekta Ukrainskogo Jazyka (Dynamics of the Hutsul Dialect of the Ukrainian Language). *Karpato-Balkanskij Dialektnij Landšaft: Jazyk i Kul'tura 2009-2011*, 2:12-31 (in Ukrainian).
- Lončarić, Celinić. 2007. Susret Slovenskih Prekmurskih i Hrvatskih Međimurskih Govora (Transitional Dialects between Slovene Prekmurjian and Croatian). *Slavistična Revija* št. 1-2 (in Serbo-Croatian).
- Lončarić, M. 1996. Kajkavsko Narječje (Kaikavian Narcissus). Zagreb, Croatia: Školska Knjiga (in Serbo-Croatian).
- Lunt, Horace. 1997. Notes on the Rusyn language of Yugoslavia and Its East Slovak origins. *International Journal of Slavic Linguistics and Poetics* XLII, pp. 44-84. Columbus, Ohio.
- Machek, Václav. 1997. Etymologický Slovník Jazyka Českého (Etymological Dictionary of the Czech Language). (In Czech).
- Majewicz, Alfred F. 1986. A New Kashubian Dictionary and the Problem of the Linguistic Status of Kashubian, in *Collectanea Linguistica in Honorem Adami Heinz*, 95-99. Warsaw: PAN/Ossolineum.
- Maksimovich, Walter. October 9, 2001. Ruska Narodna Respublika (Lemko Republic of Florynka). Lemkos.org.
- Maksimovich, Walter. February 21, 2008. The Rusyn Language: Recent Achievements and Present Challenges. Lemkos.org.
- Maksymiuk, Jan. 2014. "The Standardization of a Latin-based Orthography in Podlachian," in *Dialogue on Dialect Standardization*, edited by Carrie Dyck, Tania Granadillo, Keren Rice. Newcastle on Tyne: Cambridge Scholars Publishing.

Mandić, David. 2012. The Old Isoglosses of Southern Istria. 22, 447-460.

Mandić, D. 2013. The Accentuation of I-Verbs in the Local Dialect of Banjole. 39, 69-82.

Martin, Lucille. May-June 2013. [Languages in Peril – The Polish Connection](#), in *Parrot Time Magazine*. Scriveremo Publishing.

Matasović, Ranko. 2008. Poredbenopovijesna Gramatika Hrvatskoga Jezika (Comparative Historical Grammar of the Croatian Language). Zagreb, Croatia: Matica Hrvatska (in Serbo-Croatian).

Mezentseva, Inna. December 2014-May 2015). English professor, Belarussian and Russian speaker, BA in Education and Linguistics, Vitebsk State University, Vitebsk, Belarus. Personal communication.

Mihaljević, Milan (2002). Slavenska Poredbena Gramatika, 1. Dio, Uvod i Fonologija (Slavic Comparative Grammar, Part 1: Introduction and Phonology). Zagreb: Školska Knjiga (in Serbo-Croatian).

Moseley, Christopher, editor. 2010. [Atlas of the World's Languages in Danger](#), 3rd ed., p. 37. Paris: UNESCO Publishing.

Nábělková, M. 2007. Closely Related Languages in Contact: Czech, Slovak, "Czechoslovak". *International Journal of the Sociology of Language* 183: 53–73.

Nedelchev, Neno. 1994: Диалект на българите-католици. Северен павликянски говор (The Language of the Bulgarian Catholics: North Paulician speech.) Veliko Tarnovo (in Bulgarian).

Nikitin, A. G., I. T. Kochkin, C. M. June, C. M. Willis, I. McBain, and M. Y. Videiko. February 2009. Mitochondrial DNA Sequence Variation in Boyko, Hutsul and Lemko Populations of Carpathian Highlands. *Human Biology* 81:1, 43–58.

Nuorluoto, Juhani. 2010. Central Slovak and Kaikavian Structural Convergences: A Tentative Survey. *Journal of Slavic Languages and Literatures* 50, 37-45. Uppsala, Finland.

Pavlenko, A. 2006. Russian as a Lingua Franca. *Annual Review of Applied Linguistics* 26: 78-99.

Pavličević-Franić, Dunja and Katarina Aladrović Slovaček. 2017. Mastering Croatian Standard Language in a Plurilingual School

Environment, in Croatian Journal of Education.

Pawischitz, Sabine. 2014. Burgenland-Croatian: First Signs of Language Decay, in *Slavic Eurasian Studies* No. 26. *Slavic and German in Contact: Studies from Areal and Contrastive Linguistics*, pp. 59-80, edited by Elżbieta Kaczmarska and Motoki Nomachi. Sapporo, Japan: Hokkaido University Slavic Research Center.

Paukowits, Frank. July 31, 2014. Burgenland Bunch Newsletter. The Endangered Burgenland-Croatian Language (With Contributions from Frank Paukowits). The Burgenland Bunch Issue 245.

Petričević, Ivo. February-March 2019. Croatian speaker, Croatia. Personal communication.

Petrov, Grgo. September 2015. Kaikavian, Čakavian, and Štokavian speaker, Kaikavian and Čakavian language activist, Zagreb, Croatia. Personal communication.

Plishkova, Anna. 2008. Practical Spheres of the Rusyn Language in Slovakia. *Studio Slavica Hungarica* 53:1, pp. 95-115.

Pliško, Lina. 2009. Jugozapadni Istarski ili Štakavsko-Čakavski Dijalekt u Svjetlu Hrastina Istraživanja. *Hrvatski Dijalektološki Zbornik* 15: 111-124

Popović, Ivan. 1960. *Geschichte der Serbokroatischen Sprache (The History of the Serbo-Croatian Language)*. Wiesbaden, Germany: Otto Harrassowitz (in German).

Pourheidari, Gertraud. October 24, 2006. Acquaviva Collecroce. Department of Slavic Linguistics, Constance University, Constance, Germany.

Rač, M. and Lovrić, A. Ž. 2006. Naravoslovno Nazivlje Istarskih Polukajkavaca oko Buzeta (Narrative Terminology of Istrian Kaikavian around Buzet). *Zbornik Tjedna Kajkavske Kulture Knj. 2. Muži zagorskog srca. Zabok-Krapina* (in Serbo-Croatian).

Roos, Marti, Hans Nugteren, and Zhong Jìnwén. July 28, 2009. The Turkic Language Family.

Rudnyckyj, Jaroslav. 1965. *Ukrajins'ka Mova ta Jiji Hovory (The Ukrainian Language and Its Dialects)*. Winnipeg, Canada: Tovarystva Plekannya Ridnoji Movy (in Ukrainian).

Rusinko, Elaine. 2003. Straddling Borders: Literature and Identity

in Subcarpathian Rus'. Toronto, Canada: University of Toronto Press.

Sexton, M. T. and Shapoval, V. 2002. Geneological Classification of Languages. Novosibirsk, Russia.

Silić, Josip. 1998. Hrvatski Standardni Jezik i Hrvatska Narječja (Standard Croatian and Croatian Dialects). *KOLO, Godište VIII:4*. (in Serbo-Croatian).

Smiljanic, Rajka. 2013. Lexical, Pragmatic, and Positional Effects on Prosody in Two Dialects of Croatian and Serbian: An Acoustic Study, p. 29. New York: Routledge.

Shevelov, George. 1981. "Evolution of the Ukrainian Literary Language," in *Rethinking Ukrainian History*. Ivan Lysiak Rudnytsky and John-Paul Himka, editors, pp. 221–225. Edmonton, Canada: University of Alberta, Canadian Institute of Ukrainian Studies.

Sloboda, Marian and Brankaikec, Katja. August 21-24, 2012. "The Mutual Intelligibility of Slavic Languages and Its Potential for the Sorbian Language Revival," paper presented at Sociolinguistics Symposium 19. Berlin: Freie Universität Berlin.

Syniawa, Mirosław. 2010. *Ślabikörz niy dlō bajtli abo lekcyje ślōnskij gōdki*. Chorzów: Pro Loquela Silesiana Towarzystwo Kultuwowania i Promowania Ślōskiej Mowy.

Stadniczenko, J. 1947-48. Rok Szkolny 1947/1948 na wsi Opolskiej (School Year 1947/1948 in the Village of Opolska), p. 399 (in Polish).

Strazny, Phillip. 2013. *Encyclopedia of Linguistics*. New York: Routledge.

Sussex, Roland and Cubberley, Paul. 2011. The Slavic Languages. Cambridge, UK: Cambridge University Press.

Swarte, Femke, Stefanie Voigt, Jelena Golubovic, Charlotte Gooskens, Wilbert Heeringa, Anja Schüppert, Vincent Van Heuven, Renée Van Bezooijen, and Nanna Haug Hilton. August 23, 2012. "Mutual Intelligibility in the Germanic, Romance, and Slavic Language Area," paper presented at the 19th Sociolinguistics Symposium, Berlin.

Toops, Gary H.. March 2001. Book Review of Das Sorbische im Slawischen Kontext. Ausgewählte Studien. Schriften des Sorbischen

Instituts/Spisy Serbskeho Instituta, 24 by Heinz Schuster-Šewc. *Canadian Slavonic Papers/Revue Canadienne des Slavistes*, Vol. 43:1, pp. 164-166.

Univerza v Mariboru in Zgodovinsko Društvo Maribor (University of Maribor and the Historical Society of Maribor). 1988. *Časopis Za Zgodovino in Narodopisje (Review for History and Ethnography)* 24. Bloomington, IN: Indiana University Press (in Slovenian).

Vanko, Juraj. 1997. The Rusyn Language in Slovakia: Between a Rock and a Hard Place. *International Journal of the Sociology of Language* 183: 75-96.

Willem Vermeer. 1982. On the Principal Sources for the Study of Čakavian Dialects with Neocircumflex in Adjectives and e-Presents. *Studies in Slavic and General Linguistics* 2: 279-340.

Weigand, Gustav. 1925. "Vorwort, Zugleich Programm des Balkan-Archivs (Forward and Introduction to the Program of the Balkan Archiv)", in *Balkan-Archiv. Fortsetzung des Jahresberichtes des Instituts für Rumänische Sprache. I. Band. S. V-XV. Leipzig* (in German).

Yoshamya, Mitjel and Zyelimer. 2005. *Gan-Veyan: Neo-Liburnic Glossary, Grammar, Culture, Genome, Old-Croatian Archidioms. Monograph I. Scientific Society for Ethnogenesis Studies: Zagreb.*

Želechowskij, S. 1884. *Заметки о русских говорах Седлецкой губернии (Notes on the Russian Dialects of Sedletska Province)*. *Русский Филологический Вестник*, t.XI, p. 258 (in Russian).

Zušťák, Michal. April-July 2015. Slovak speaker and law student, Bratislava, Slovakia. Personal communication.