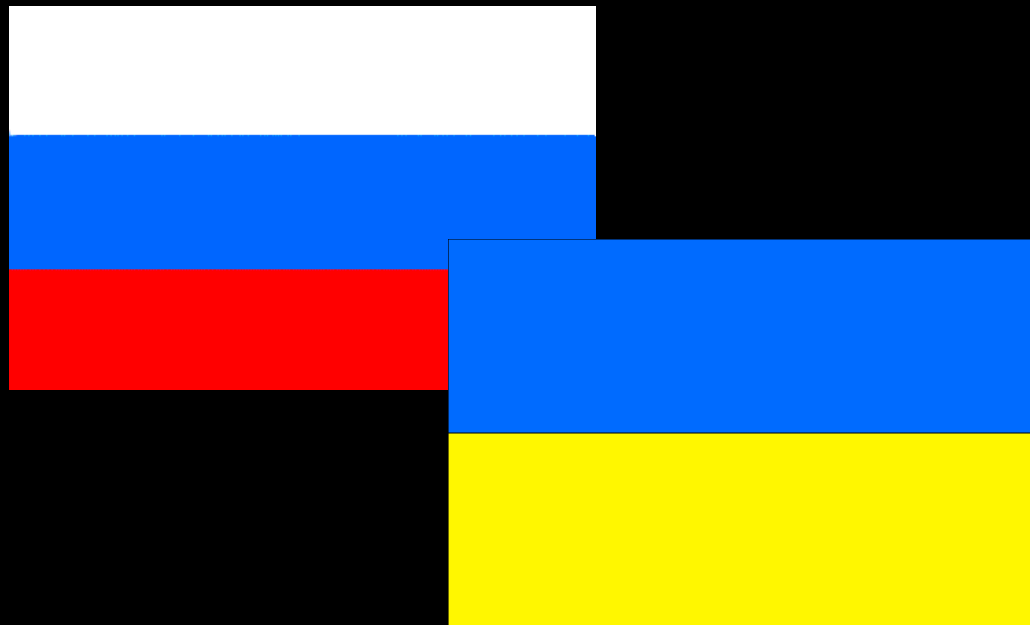


Der Stellenwert des Ukrainischen und die Idee eines russisch-ukrainischen Parallelkorpus



Дякую Тобі,
Боже,



що я
НЕ МОСКАЛЬ!

ДЯКУЮ ТОБІ,
БОЖЕ,



ЩО Я
НЕ МОСКАЛЬ!



Білорусь

Росія

Польща

Луцьк

Рівне

Чернігів

Суми

Житомир

Харків

Львів

Тернопіль

Хмельницький

Вінниця

Черкаси

Полтава

Ужгород

Івано-Франківськ

Чернівці

Кіровоград

Дніпропетровськ

Донецьк

Запоріжжя

Луганськ

Молдова

Румунія

Миколаїв

Одеса

Херсон

Азовське море

УКРАЇНА

обласний центр



область



ріки

0 50 100 150 200 км

Чорне море

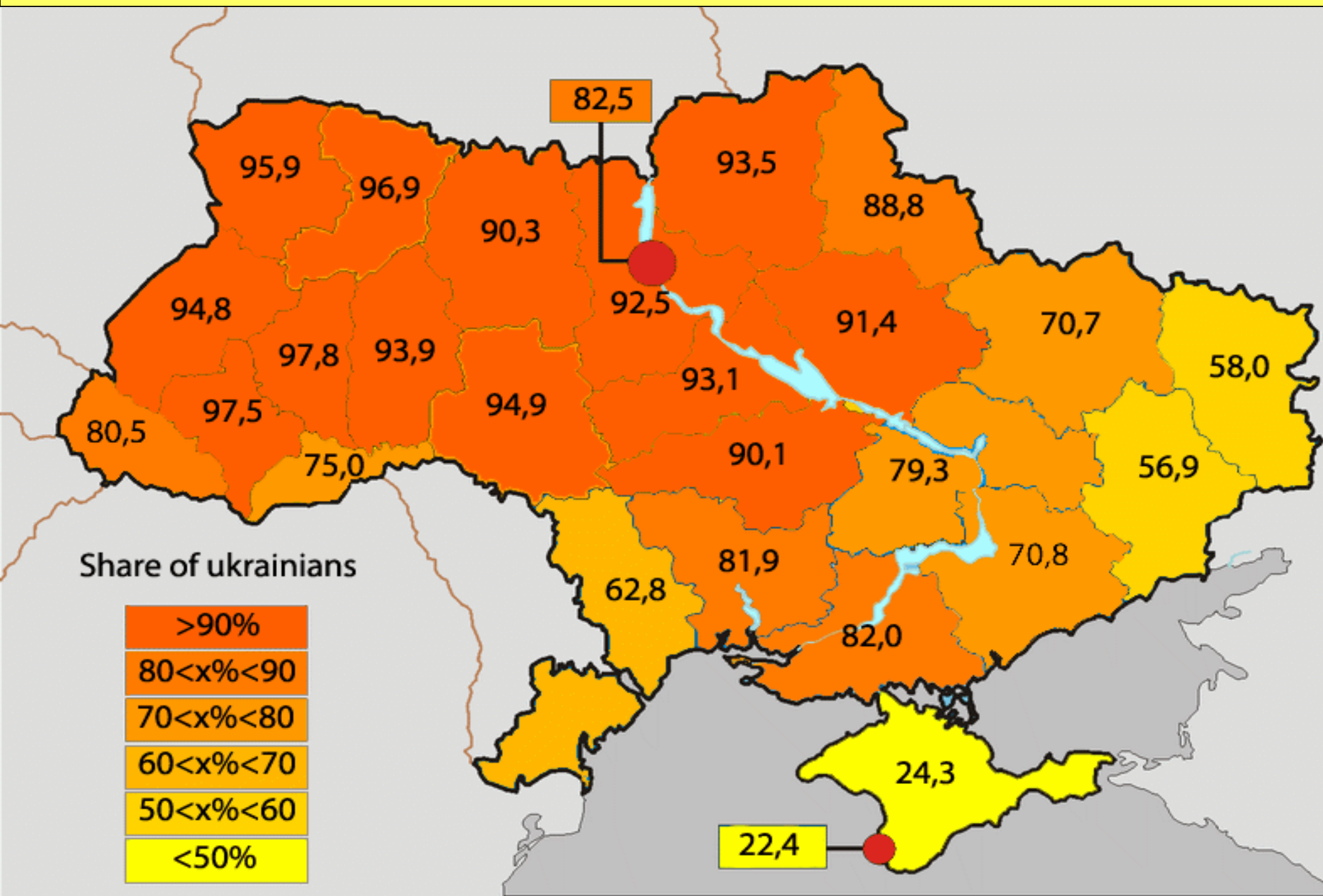
Сімферополь

Севастополь

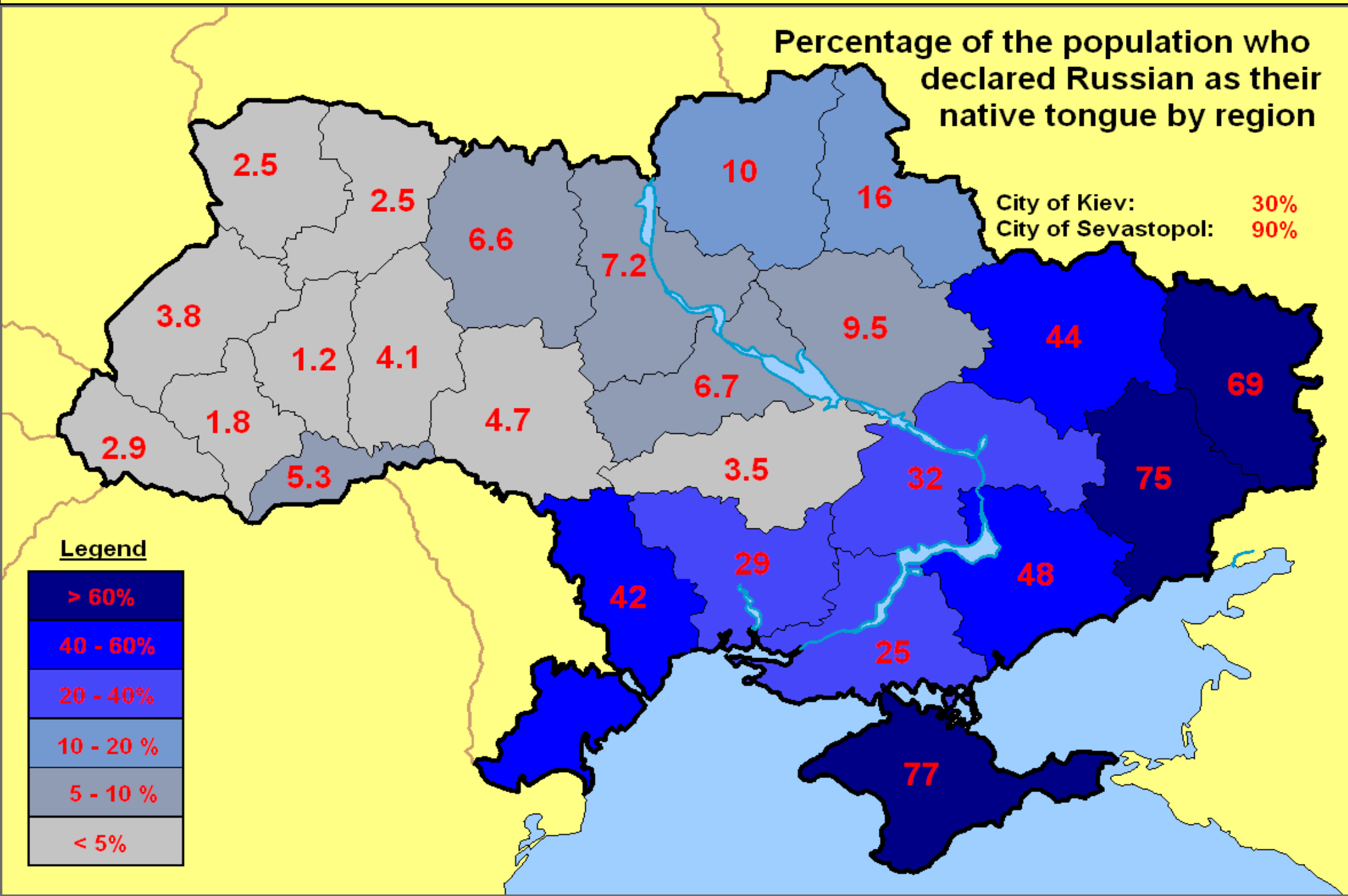
OFFICIAL BREAKDOWN OF UKRAINE'S PRESIDENTIAL VOTE



Offizielle Karte der Sprecher des Ukrainischen

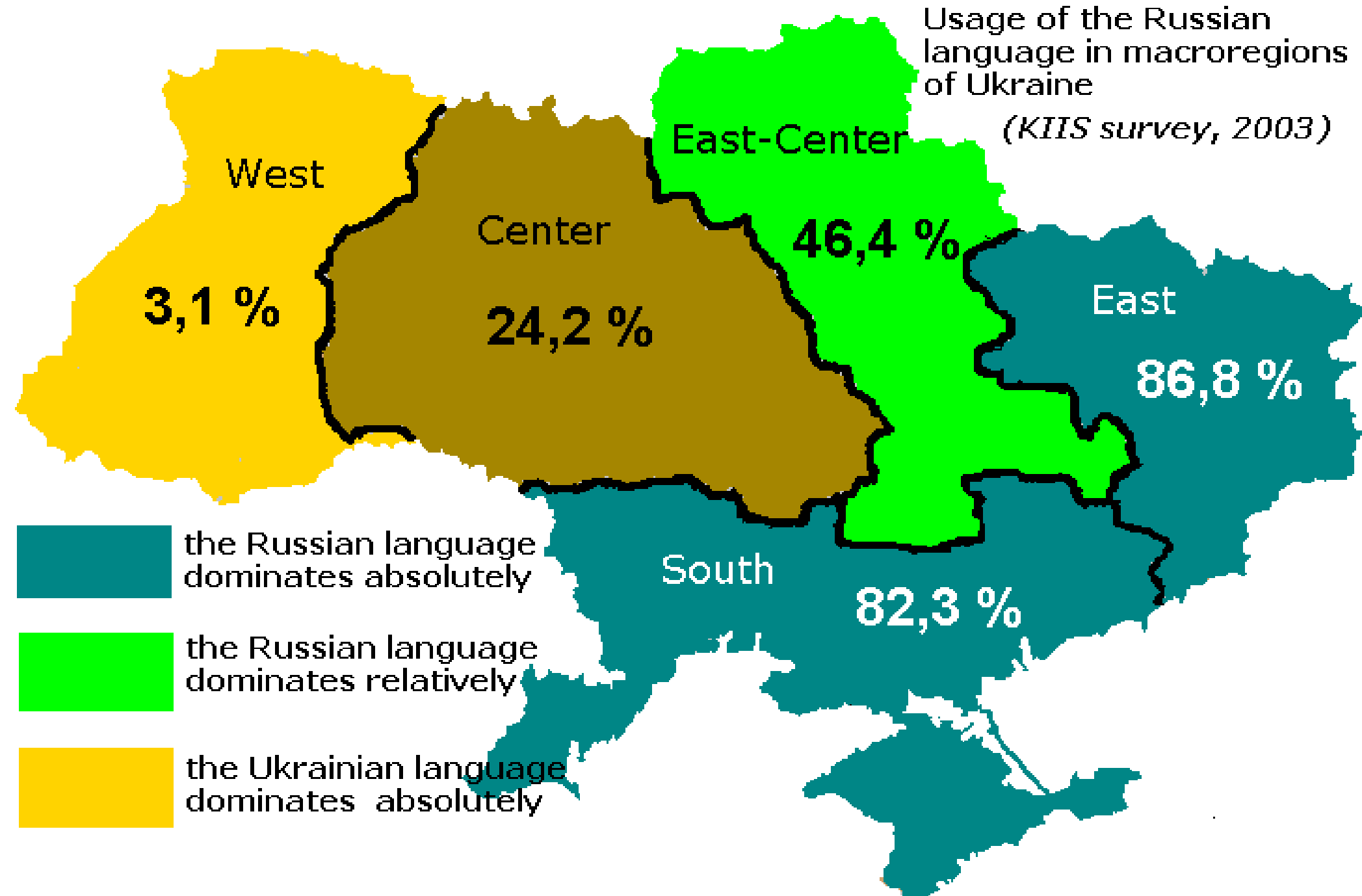


Offizielle Karte der Sprecher des Russischen

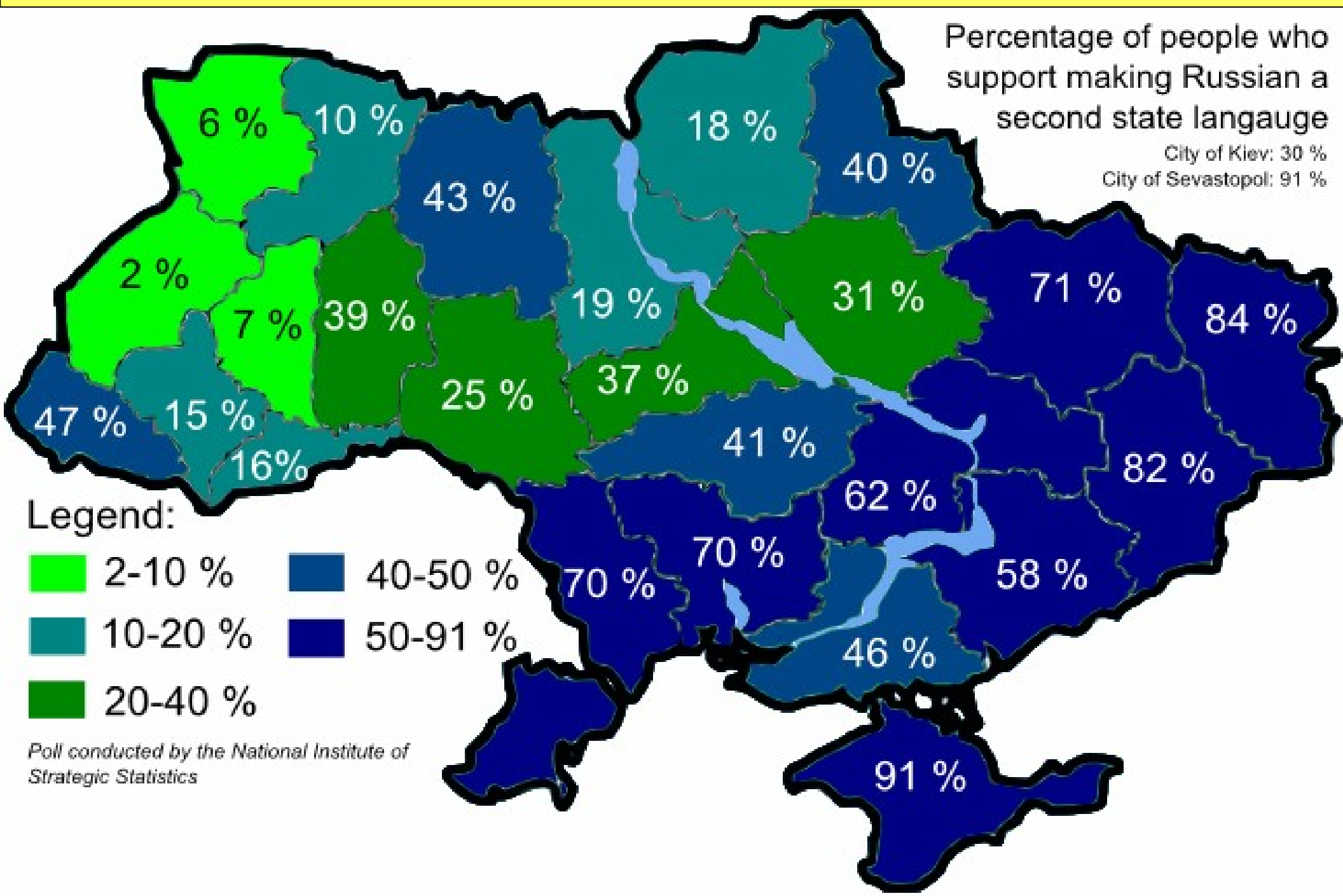


Russisch in der Ukraine (nach Makroregionen)

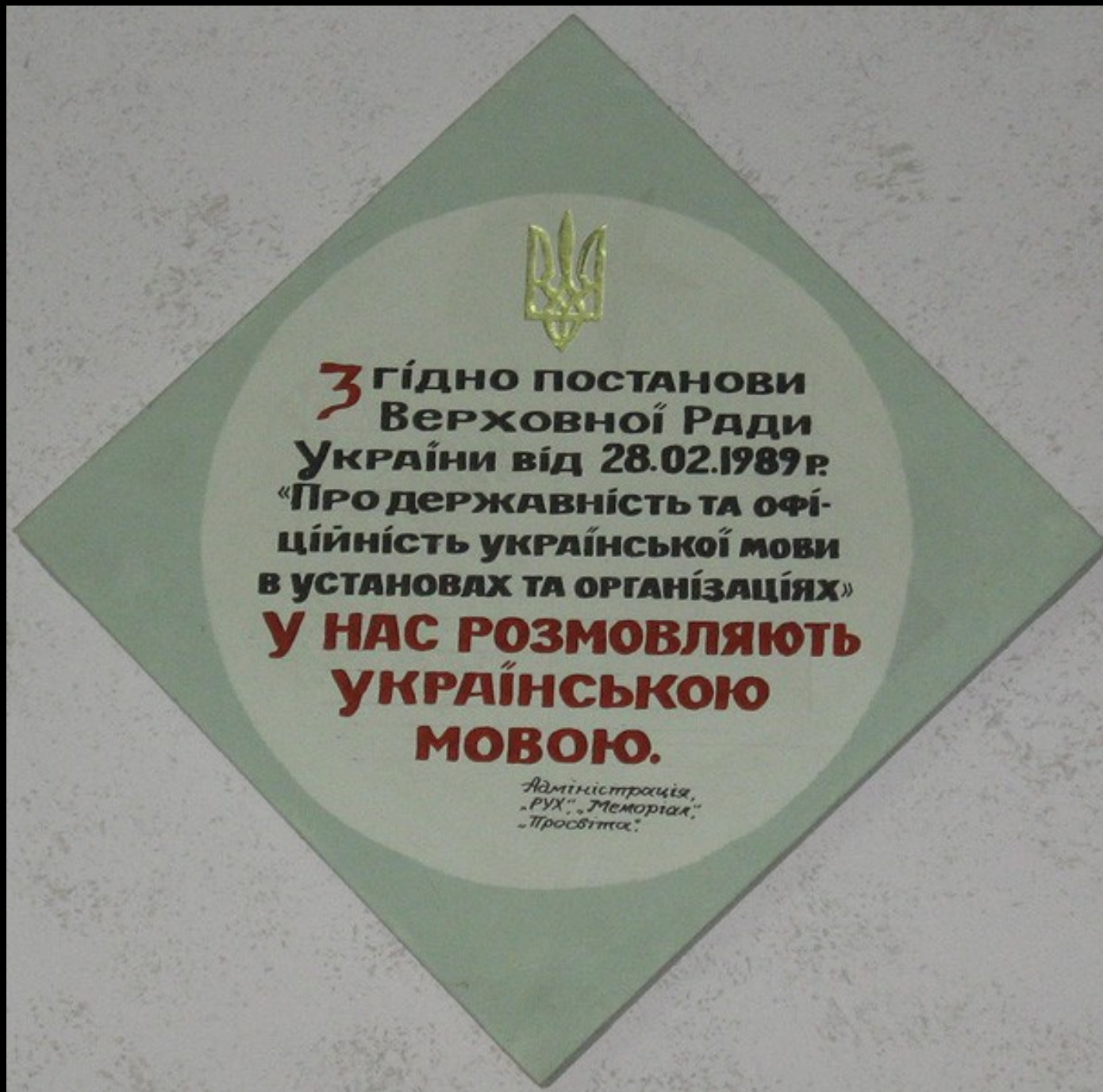
Usage of the Russian language in macroregions of Ukraine
(KIIS survey, 2003)



Russisch als zweite Amtssprache



Українізація (Ukrainisierung)



Parallelkorpus

Bitext

im Idealfall 1:1-Übersetzung
(nahezu unmöglich)

Beispiel

uk: Україна	є	гарна	країна
ru: Украина	-	красива	страна

БаДаРУЯ

База ДАнных Русского и Украинского Языков



База ДАнних Російської та Української Мов

БаДаРУМ

Korpusinformation

Name: BaDaRUJa/BaDaRUM

Art: Experimenteller Parallelkorpus

Sprachen: Russisch, Ukrainisch

Textbasis: Vorwort, Werbung, Debian

Token: ca. 8000

Zeit: Nov 2006 – Feb 2007

Code: XCES (X Corpus Encoding Standard)

XML-Code BaDaRUJa/BaDaRUM

```
6884 </w id="w186.1.11" type="S" lemma="сайт">сайт</w>
6885 <w id="w186.1.12" type="S" lemma="сайт">сайті</w>
6886 <w id="w186.1.13" type="PUNCT" lemma=".">.</w>
6887 </s>
6888 </p>
6889 <head id="187">
6890 <s id="s187.1">
6891 <w id="w187.1.1" type="A" lemma="який">Який</w>
6892 <w id="w187.1.2" type="S" lemma="формат">формат</w>
6893 <w id="w187.1.3" type="A" lemma="двійковий">двійкових</w>
6894 <w id="w187.1.4" type="S" lemma="пакунок">пакунків</w>
6895 <w id="w187.1.5" type="COMP" lemma="Debian">Debian</w>
6896 <w id="w187.1.6" type="SENT" lemma="?">?</w>
6897 </s>
6898 </head>
6899 <p id="188">
6900 <s id="s188.1">
6901 <w id="w188.1.1" type="S" lemma="пакунок">Пакунок</w>
6902 <w id="w188.1.2" type="PUNCT" lemma=",">,</w>
6903 <w id="w188.1.3" type="CONJ" lemma="або">або</w>
6904 <w id="w188.1.4" type="PART" lemma="ж">ж</w>
6905 <w id="w188.1.5" type="A" lemma="архівний">архівний</w>
6906 <w id="w188.1.6" type="S" lemma="файл">файл</w>
6907 <w id="w188.1.7" type="COMP" lemma="Debian">Debian</w>
```

ЛиКоРУЯ

Лингвистический Корпус Русского и Украинского Языков



Лінгвістичний Корпус Російської та Української Мов

ЛіКоРУМ

Korpusinformation

Name: LiKoRUJa/LiKoRUM

Art: GRALIS-Parallelkorpus

Sprachen: Russisch, Ukrainisch

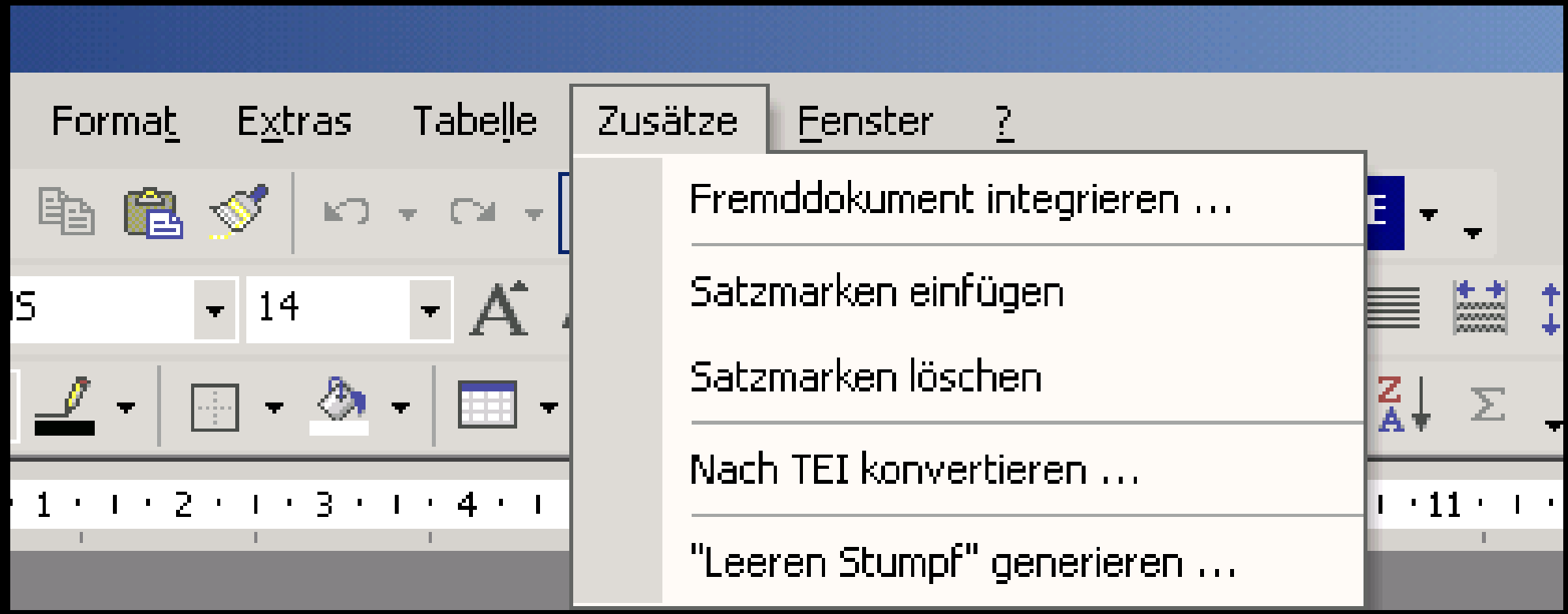
Textbasis: open source-Dokumentationen

Token: noch nicht bekannt

Zeit: Apr 2008 – Jan 2009

Code: TEI (Text Encoding Initiative)

Makros in MS-Word



WYSIWYG-Methode

Що таке Debian?

Debian — це вільна операційна система (ОС) для вашого комп'ютера.] Операційна система — це набір із базових програм та інструментів, що потрібні для роботи вашого комп'ютера.] Debian використовує ядро Linux але більшість основних програм взято з проекту GNU, тому система називається GNU/Linux.]

Debian GNU/Linux — це не просто операційна система:] в неї входять понад 18733 пакунки скомпільованого програмного забезпечення, які легко можуть бути встановлені на вашому комп'ютері.]

TEI-Dokument

```
29 </teiHeader>
30 <text>
31   <body>
32     <div1 type="">
33       <head lang="uk">Що таке Debian?</head>
34       <p n="0">
35         <s lang="uk">Debian — це вільна операційна система (ОС) для вашого комп'ютера.</s>
36         <s lang="uk">Операційна система — це набір із базових програм та інструментів, що
37           потрібні для роботи вашого комп'ютера.</s>
38         <s lang="uk">Debian використовує ядро Linux але більшість основних програм взято з
39           проекту GNU, тому система називається GNU/Linux. </s>
40         <s lang="uk"/>
41       </p>
42       <p n="1">
43         <s lang="uk">Debian GNU/Linux — це не просто операційна система: </s>
44         <s lang="uk">в неї входять понад 18733 пакунки скомпільованого програмного
45           забезпечення, які легко можуть бути встановлені на вашому комп'ютері. </s>
46         <s lang="uk"/>
47       </p>
48       <p n="2">
49         <s lang="uk"/>
50       </p>
51     </div1>
52   </body>
53 </text>
54 </TEI.2>
```

de

en

uk

ru

Danke
für
Ihre
Aufmerksamkeit
!

Thank
you
for
your
attention
!

Дякую
за
Вашу
увагу
!

Спасибо
за
Ваше
внимание
!