

# Global Health Applications

Creating Multilingual Reports with the SAS Unicode Server

Sharon Coleman MS, MPH

December 3, 2015



# Introduction

- Global public health initiatives are critical to improving the health of populations worldwide
- Research that incorporates data from various countries with different languages and scripts can be challenging
- The SAS Unicode server allows encoding of characters from multiple languages
- Bridging the language divide in research

# What is the Unicode Transformation Format?

- A SAS session that uses UTF-8 : universal coding
- ASCII (American Standard Code for Information Interchange) codes represent text in computers and was the most common character encoding until 2007. Then UTF-8 surpassed ASCII
- ASCII only allows English and common punctuation but isn't useful for non-Latin characters  
128 ASCII characters vs. over a million code points!
- Code points are used with UTF-8 (**10101100** for "€" )

# Installing the Unicode server

- On Windows, you will find the additional language option on the SAS programs menu

**Programs ► ► SAS**

**► SAS 9.3 (English)**

**► SAS 9.3 (Additional Languages)**

**► SAS 9.3 (Unicode support)**

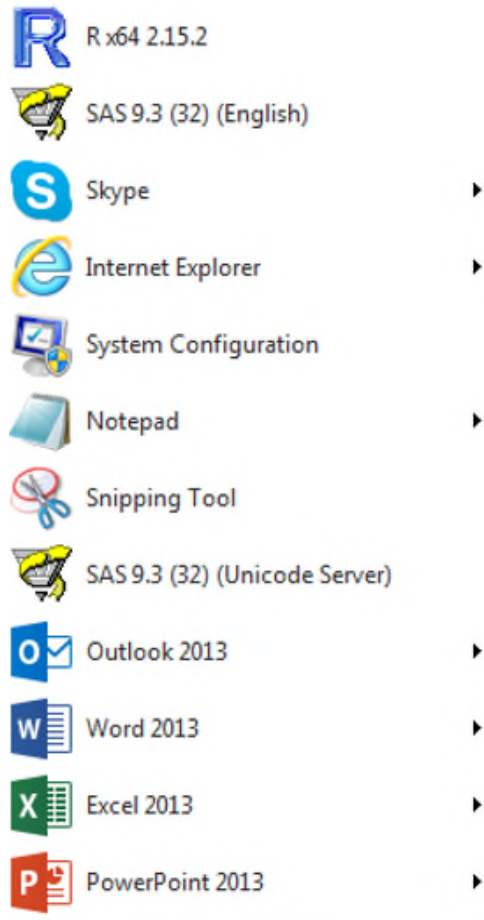
\*Please see contact SAS support for specific questions

<http://support.sas.com/techsup/technote/ts801.pdf> (p 6-7)

# Multiple languages can be used in one session

|    | name      | first     | street               | zip      | city        | country |
|----|-----------|-----------|----------------------|----------|-------------|---------|
| 60 | Côté      | Frédéric  | 2, rue descente...   | 36000    | Châteauroux | France  |
| 61 | Boucher   | Corinne   | 21 place du Pant...  | 75009    | Paris       | France  |
| 62 | Fournier  | Étienne   | 91, rue Victor Hu... | 71000    | Mâcon       | France  |
| 63 | Cotée     | Madeleine | 691, avenue Fré...   | 30034    | Nîmes       | France  |
| 64 | Legrand   | Clare     | 62 rue de la Biet... | 45000    | Orléans     | France  |
| 65 | Dubois    | Benoît    | 40 Rue Charles...    | 91330    | Yerres      | France  |
| 66 | Thibeault | Georges   | 47 rue Alexandre...  | 91000    | Évry        | France  |
| 67 | Martin    | Désirée   | 90, rue Grisolle     | 83600    | Fréjus      | France  |
| 68 | Vaudron   | Sébastien | 2 promenade de...    | 13008    | Marseille   | France  |
| 69 | Girard    | Régine    | 4 bis rue Profess... | 69008    | Lyon        | France  |
| 70 | Yılmaz    | Ekrem     | Tepebaşı Bulvarı...  | 80050    | İstanbul    | Türkiye |
| 71 | Kaya      | Erkan     | Akay Caddesi N...    | 06640    | Ankara      | Türkiye |
| 72 | Demir     | Filiz     | Fethi Cadessi No...  | 33212    | Izmir       | Türkiye |
| 73 | Şahin     | Ali       | Sanayi Cad. No...    | 23300    | Elazığ      | Türkiye |
| 74 | Çelik     | Nur       | Kazımiye Mah. O...   | 59860    | Çorlu       | Türkiye |
| 75 | 佐藤        | 明子        | 東京都渋谷区...            | 〒150-000 | 東京          | 日本      |
| 76 | 鈴木        | 幹夫        | 神奈川県横浜...            | 〒221-083 | 横浜市         | 日本      |
| 77 | 田中        | あずみ       | 大阪府大阪市...            | 〒534-001 | 大阪市         | 日本      |
| 78 | 山本        | 麻美        | 厚木市寿町3-7...          | 〒243-000 | 厚木市         | 日本      |

# Start a SAS session that uses the UTF-8





# Goal: Create a report of follow up study visits

Colleagues are in St. Petersburg  
Russia

Assistance needed in seeing  
follow up visits to implement  
study reminder calls



# Example: Writing a report using Cyrillic script

| subjectid | firstname | lastname   | hphone1         | cphone1         | visit     | scheddated |
|-----------|-----------|------------|-----------------|-----------------|-----------|------------|
| 1002      | Олег      | Чаплыгина  | 8-812-750-62-66 | 8-812-750-62-66 | 6 месяцев | 3/12/2015  |
| 1004      | Роман     | Викторович | 8-812-750-62-66 | 8-812-755-66-66 | 7 месяцев | 3/12/2015  |
| 1005      | Алексей   | Намик-оглы | 8-812-750-62-66 | 8-813-500-62-66 | 8 месяцев | 3/12/2015  |

```
data russia;  
set work.russia;  
format scheddated mmddyy8.;  
run;
```



# Example: Writing a report using Cyrillic script

```
ods escapechar = '^' ;
options orientation = landscape nodate nonumber ;
ods noproctitle;
ods pdf file="C:\basug\scheduled_visits.pdf"
startpage=on startpage=now;
title1 bold height=11pt "Россия:Назначенные Отдаленные Интервью";
title2 bold height=9pt "Scheduled Follow-Up Visits";
title3 bold height=10pt " &dtnull13. ";
□ proc report data=russia nowd headline headskip box STYLE(header)=[font_weight=bold background=CXAEADD9
vjust = center] missing;

column subjectid firstname lastname hphone1 cphone1 visit scheddate ;

define subjectid/center "Номер участника" style=[cellwidth=0.7in];
define firstname/left "Имя" style=[cellwidth=1.2in];
define lastname/left "Фамилия" style=[cellwidth=1.2in];
define hphone1/left "Домашний Телефон" style=[cellwidth=1.2in];
define cphone1/left "Мобильный Телефон" style=[cellwidth=1.2in];
define visit/left "Вид Визита" style=[cellwidth=0.8in];
define scheddate/left "Назначенная дата" style=[cellwidth=0.8in];

run;

ods pdf close;
```

# Warning!

- When you open SAS with Unicode Support, you will see the following in the log

WARNING: Display of UTF8 encoded data is not fully supported by the SAS Display Manager System.

You could use SAS Enterprise Guide or use an ODS output option

# Output in Windows Display Manager

|      | □□□□    | □□□□□□□□    | □□□□□□□□ □□□□□□ |
|------|---------|-------------|-----------------|
| 1002 | □□□□    | □□□□□□□□    | 8-812-750-62-66 |
| 1004 | □□□□□   | □□□□□□□□□□  | 8-812-750-62-66 |
| 1005 | □□□□□□□ | □□□□□□-□□□□ | 8-812-750-62-66 |

# Report that is translated for the Russian team

***Россия: Назначенные Отдаленные Интервью***  
***Scheduled Follow-Up Visits***  
***3 December 2015***

| Номер участника | Имя     | Фамилия    | Домашний Телефон | Мобильный Телефон | Вид Визита | Назначенная дата |
|-----------------|---------|------------|------------------|-------------------|------------|------------------|
| 1002            | Олег    | Чаплыгина  | 8-812-750-62-66  | 8-812-750-62-66   | 6 месяцев  | 03/12/15         |
| 1004            | Роман   | Викторович | 8-812-750-62-66  | 8-812-755-66-66   | 7 месяцев  | 03/12/15         |
| 1005            | Алексей | Намик-оглы | 8-812-750-62-66  | 8-813-500-62-66   | 8 месяцев  | 03/12/15         |

# Using National Language Support (NLS)

- Refer to —Values for the LOCALE= System Option SAS

<http://support.sas.com/documentation/cdl/en/nlsref/63072/PDF/default/nlsref.pdf>

# Locale sensitive formatting

- `option locale=ru_RU;`
- `data russia2;`
- `set work.russia2;`
- `format Compensation NLMNYI.;`
- `run;`
- `proc print;var subjectid lastname Compensation;`  
`run;`



# Currency specific to the Country

| Наблюдения | subjectid | lastname   | Compensation |
|------------|-----------|------------|--------------|
| 1          | 1002      | Чаплыгина  | 1 000RUB     |
| 2          | 1004      | Викторович | 1 500RUB     |
| 3          | 1005      | Намик-оглы | 2 000RUB     |

# Goal: Create a report of Chinese SMS messages

Colleagues are in China

Patient navigators will text a message in the participants preferred language



# Data from an Excel spreadsheet

| Subjectid | cellphone  | SMS Text |
|-----------|------------|----------|
| 1001      | 0839947259 | 安排复诊时间   |
| 1002      | 0839947259 | 检查结果     |
| 1003      | 0839947259 | 检查结果     |

```
proc import out= work.china
datafile= "C:\basug\china_report.xlsx"
dbms=excel replace;
getnames=yes;
mixed=no;
scantext=yes;
usedate=yes;
scantime=yes;
run;
```

```
ods escapechar = '^' ;
options orientation = landscape nodate nonumber ;
ods noproctitle;
ods rtf file="C:\basug\china_sms.doc"
startpage=on startpage=now;
title1 bold height=11pt "Preferred Language is Chinese";
title2 bold height=9pt "SMS Message";
title3 bold height=10pt " 3 December 2015 ";
footnote1 " Schedule Medical f/u visit";
footnote2 "check lab results";
```

```
proc report data=china nowd headline headskip box STYLE(header)=[font_weight=bold background=CXAEADD
vjust = center] missing;
column subjectid cellphone SMS_Text ;
```

```
define subjectid/center      "SubjectID"          style=[cellwidth=0.8in];
define cellphone/left       "CellPhone"           style=[cellwidth=1.2in];
define SMS_Text/left        "SMS Text"             style=[cellwidth=1.5in];
run;
```

```
ods rtf close;
```

# SMS text to be sent in Chinese characters

*Prefered Language is Chinese*  
*SMS Message*  
*3 December 2015*

| <u>SubjectID</u> | <u>CellPhone</u> | SMS Text      |
|------------------|------------------|---------------|
| 1001             | 0839947259       | <u>安排复诊时间</u> |
| 1002             | 0839947259       | <u>检查结果</u>   |
| 1003             | 0839947259       | <u>检查结果</u>   |

# Reference

- An excellent reference :

For more information about the concept of the SAS Unicode Server, see the technical Paper, —Processing Multilingual Data with the SAS® 9.2 Unicode Server.

# Conclusion

- With the SAS Unicode server , you can meet global public health and research needs in a single SAS session!

Thank you: спасибо

Sharon Coleman  
sharcole@bu.edu

