Different languages; different protocols

2

- representation and annotation of English in transcripts of conversations in six Heritage Languages (Cantonese, Faetar, Italian/Calabrese, Korean, Russian, and Ukrainian)

- Only Faetar is translated to English.

- Methods differ by research team

- Best option for standardizing?

naomi.nagy@utoronto.ca

Russian

3

- 3-letter language tag "ENG" (or another language) introduces any non-Russian word/phrase, which is bracketed
- English words with Russian morphemes are transcribed as Russian
  - "Whether we use English spelling or transliterate the utterance depends largely on how the speaker says it, whether they use English-like or Russian-like pronunciation.
    - Context - are they deliberately using an English word? struggling to find a Russian equivalent?
    - Abbreviations like "UofT" will be written in English because they are abbreviations of English words. It's possible to come up with abbreviations of Russian words for the same concepts." [NL]
- transcription is transliteratable with Comrie & Corbett's (2002) system, at http://www.translit.ru/

naomi.nagy@utoronto.ca

Examples from Russian

4

- Aga, ja prepodavala francuzskij v [ENG: UofT], jeto bylo vsjo [ENG: part-time].

Yes, I taught French at UofT, it was all part-time.
[R1F55B_IV_PR.eaf, 2:38]

- Tam oni ochen' mnogo tam [ENG: fundraising] i tam raznyx vesjolyx veshhej.

There they do a lot of fundraising and various fun things.
[R2F12A_IV_PR.eaf, 0:24]

  - Regular expression searchable: **ENG:**

naomi.nagy@utoronto.ca

5

English words with Russian morphemes

- da, vam poslajsat' kolbasku ili pisikom, da.

Yes, would you like your kielbasa sliced or in one piece, yes. [R1M56A_IV_PR.eaf, 0:29:23]

poslajsat	pisikom
po+*slice*+at'	*piece*+ik+o
"to slice"	"in one piece"

- Not regular expression searchable

6

Ukrainian

- transcribe English words with capital letters

- "If a word exists in both language, then I will listen closely to the phonology and transcribe it accordingly.

- If they pronounce an English word with a Ukrainian accent then I will transcribe it in Ukrainian, but I will make a note in the notes tier."

Slide 7

Example in ELAN - UKR
(translation added)

7


```
                        Elan – U3M41A_IV_ENG.eaf
File  Edit  Annotation  Tier  Type  Search  View  Options  Window  Help
          Grid   Text   Subtitles   Audio Recognizer   Metadata   Controls

   U3M41A

> Nr   Annotation                                                                    Begin Time    End Time     Duration
  324  che- ja dumajo shho pomohlo bo ja serjozno vs'o brav I MEAN                    00:31:16.790  00:31:22.290  00:00:05.500
  325  ale ja ne dumaju shho vono by taku velyku riznycju zrobylo jakshho by vin to ne robyv  00:31:23.270  00:31:30.030  00:00:06.760
  326  ale ven mene pxav bil'she nizh moju sestru                                     00:31:32.270  00:31:35.020  00:00:02.750
  327  UH YEAH til'ky tak SUBCONSCIOUSLY vin vin je duzhe spravedlyvyj UM v kozhnomu sensi teper  00:31:40.160  00:31:48.950  00:00:08.790
  328  ale ja dumaju v tomu chasi vin mav- EXPECTATIONS buly inakshi                  00:31:49.110  00:31:53.640  00:00:04.530
  329  to tak jak shho do matymatyky i take vo vs'o to by malo buly moji rechi RIGHT?  00:31:56.070  00:32:01.660  00:00:05.590

00:31:40.160                              Selection: 00:31:40.160 - 00:31:48.950  8790
```

U3M41A [539]: UH YEAH til'ky tak SUBCONSCIOUSLY vin vin je duzhe spravedlyvyj UM v kozhnomu sensi teper

ENG TRANS [6]: Uh yeah just subconsciously he he is very fair um in every sense now

IVer 1 [193]

IVer 2 [138]

naomi.nagy@utoronto.ca

Slide 8

8

Examples from Ukrainian

- UH YEAH til'ky tak SUBCONSCIOUSLY vin vin je duzhe spravedlyvyj UM v kozhnomu sensi teper

Uh yeah just subconsciously he he is very fair um in every sense now. [U3M41A_IV.eaf, 31:30]

- vin nazyvajet'sja ATTILIO ja joho nazyvaju ARISTOTLE

He is called Attilio, I call him Aristotle. [U2F60A_IV.eaf, 48:38]

- chasamy my jidemo do Fljorydy na MARCH BREAK i todi my USUALLY idemo do des' na lito SO

Sometimes we go to Florida for March Break and then we usually go to somewhere for the summer so [U3F13A_IV.eaf, 9:37]

- Regular expression searchable: **[ABC] or Notes tier**

naomi.nagy@utoronto.ca

9

Cantonese

- Current system: use jutping romanization
  - every Cantonese word has a number indicating tone as the final character
  - there are no tone markings on English words
  - Mandarin borrowings aren't distinguished

- Possible future system: transcribing characters
  - Cantonese and English will be even more distinct
  - Mandarin borrowings would be less obvious

naomi.nagy@utoronto.ca

10

Examples from Cantonese

- seng4 jat6, ngo5 dei6 seng4 jat6 heoi3 pet store go2 zan4 si4 le3 keoi5 hai6

When we go to the pet store she always goes -- [C2F16A_IV.eaf 29:29]

Usually like mou2 gam3 je6 la1

usually, like, not very late [C2F16A_IV.eaf 6:55]

- zing3 fu5 le1 zau6 jau5 jat1 di1 giu3 zou6 housing scheme bei5 ni1 di1

The government has something known as housing scheme to provide-- [C1M61A_IV.eaf, 4:20]

- Regular expression searchable: **[a-zA-Z]\s** (= letter followed by white space)

naomi.nagy@utoronto.ca

Korean

11

- code-switching is marked in transcriptions by use of English rather than Korean characters

- column in speaker catalog to identify speakers who code-switch to English:
  - ❑ Y (yes)
  - ❑ S (sometimes)
  - ❑ N (no)
    - (C-S was studied in the Korean corpus by a student (Chung 2010)
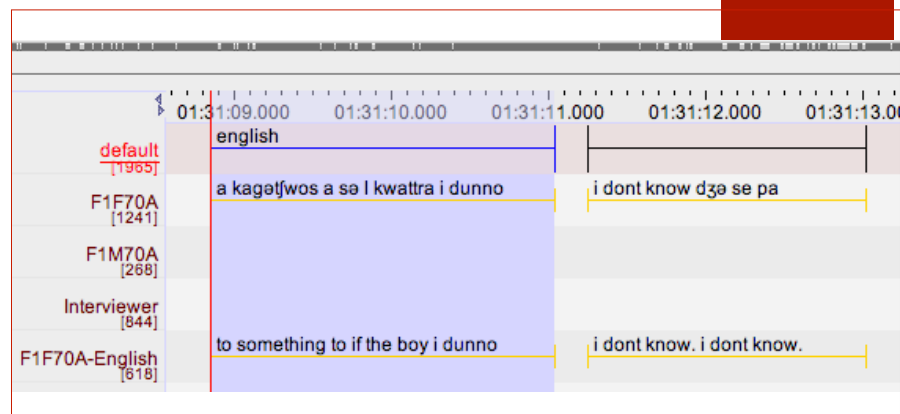
naomi.nagy@utoronto.ca

Examples from Korean

12

- 그래서 꼭 기회가 되면 , 한국에 나의 누이가 사니까 my sister lives still 그러니까 애들을 , 좀 여름 방학 때는 좀 한달이라도 좀 보내고 좀 크면 […]✝
- *So if there was an opportunity, my sister lives in Korea—my sister lives [there] still so I would like to send my kids there for a month during the summer when they get older…* [K1M45A_IV.eaf 16:16]

- 그레도 저도 이제 you know Five-six, five years 결혼 했으니깐 재 맛도 조금 틀여졌어요 my wife으로 Because she does most of the cooking.

- *Even now, I've been married for you know five-six, five years, my taste has changed a bit—my wife—Because she does most of the cooking.* [K2M34A_IV.eaf 3:05]
- Regular expression searchable: **[a-zA-Z]**

naomi.nagy@utoronto.ca

Faetar

13

- Probably not regular expression searchable: **non-IPA ?**

but "English" is noted in a separate tier or "(ITALIAN)" follows transcription

naomi.nagy@utoronto.ca

Examples from Faetar

14

sta laʃandə i pjatti nu sɛmpr doing the dishes ma

They are washing the dishes, no, always doing the dishes, but… [F1F71A_IV.eaf, 1:23:17]

- ɪn toskan i kiamuntə lə i lamponi
- *in Tuscan they call the "the raspberries" (ITALIAN)* [F1M75A&family_IV_part1.eaf, 29:26]

naomi.nagy@utoronto.ca

Italian

15

- mio padre mi aveva comprato, come si dice, una "top" you know, uno di quelli tops che si girano

My father bought me, how do you say, a "top," you know, one of those tops that spins [I2F53A_IV.eaf, 02:34]

- di solito, I ignore them (laughs).

Usually, I ignore them. [I2F53A_IV.eaf, 2:59]

- well George and I abitamo, a abitamo proprio a Avenue Rd e Eglinton

Well, George and I live, live right at Avenue Road and Eglinton. [I2F32A_IV_PT2.eaf, 2:20-2:25]

Not usually marked, not searchable

naomi.nagy@utoronto.ca

References

16

- Chung, Sheila. 2010. Code-switching as a means of cultural identity among Koreans in Toronto. TULCON '10, Cornell Undergraduate Linguistics Colloquium 2010.

- Comrie, Bernard & Greville Corbett. 2002. *The Slavonic Languages*. London & New York: Routledge. 827, 832-833.

- Farley, C. & D. Lister. 2007. Greater Toronto's language quilt. *Toronto Star*. Dec. 30, 2007.

- Keefe, S. & A. Padilla. 1987. *Chicano Ethnicity*. Albuquerque, NM: UNM Press.

- Nagy, N. 2009. Heritage Language Variation and Change. http://projects.chass.utoronto.ca/ngn/HLVC.

naomi.nagy@utoronto.ca