

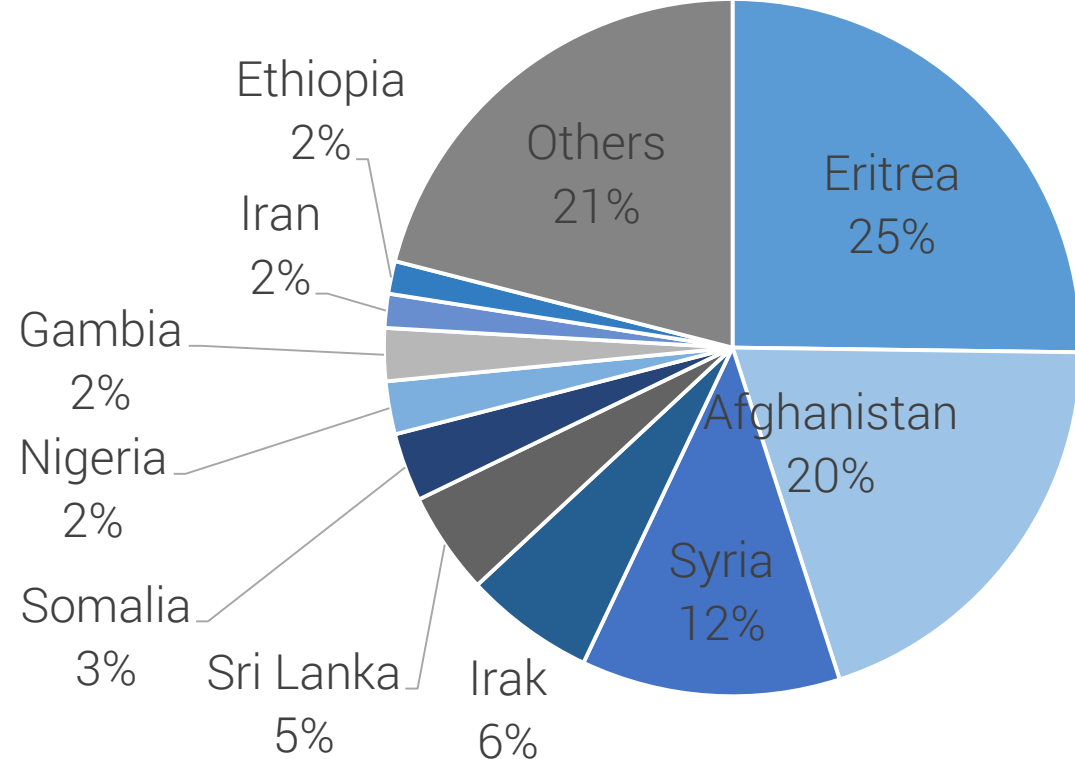
BabelDr: A Web Platform for Rapid Construction of Phrasebook-Style Medical Speech Translation Applications

Pierrette Bouillon (FTI), Hervé Spechbach (HUG), Sophie Durieux-Paillard (HUG), Johanna Gerlach (FTI), Sonia Halimi (FTI), Patricia Hudelson (HUG), Manny Rayner (FTI), Irène Strasly (FTI), Nikos Tsourakis (FTI)

Context

52% of patients at HUG (Geneva University Hospitals) are foreign ; 12% don't speak French

2015 Asylum applications in Switzerland (total: 39'523)



Source: State Secretariat for Migration (SEM), 2016, www.sem.admin.ch

Google translate is not a solution:

- not reliable and precise enough
- most important languages for HUG not covered (Tigrinya, Romani, etc.)

Solution : **BabelDr**

- Common project UNIGE/FTI-HUG
- a web-based platform for building medium-size speech translators in medical domains
- Reliable, easy to add new languages, easy to deploy

Financed by: «Fondation privée des Hôpitaux universitaires de Genève» and Innogap/Mimosa grants

The application

Examination-tool for medical subdomains : semantic coverage consists of a prespecified set of questions, but users can use a wide variety of surface forms when speaking to the system

Translation steps:

- Rule-based **recognition** of the source question (Nuance)
 - Est-ce que vous avez de la température ?
- Rule-based **translation** of the source to source canonical with synchronized CFG
 - Avez-vous de la fièvre ?
- Human translation** from source canonical to target sentence
 - ረስኒ አለካዶ?

→ Data will be used to build a robust recognizer and a classifier

Development platform and steps

Developed with Regulus Lite

- The set of canonical questions is defined by HUG
- The questions are transformed in synchronized CFG by experts (UMLS, etc.)


```
Utterance
Source ?$votre_douleur $irradie_elle $$vers_organe
Source ?(est-ce (quelqu')) $votre_douleur $irradie $$vers_organe
Source ?$vos_douleurs $irradient_elles $$vers_organe
Source ?(est-ce (quelqu')) $vos_douleurs $irradient $$vers_organe
Source $$vers_organe
Source ($savez_vous | $ça_fait | $ressentez_vous | $souffrez_vous de)
$mal_de_ventre $$vers_organe
Target/french la douleur au ventre irradie-t-elle $$vers_organe ?
EndUtterance
```
- The canonical questions are translated by translators


```
Utterance
Target/french la douleur au ventre irradie-t-elle $$vers_organe:1 ?
Target/tigrinya ለተ ናይ ከብዲ ቃንዛ ፋሕዶ ይብል ናብ $$vers_organe:1 ?
EndUtterance
```
- The system is compiled on the web

Content

- 3000 questions collected for now
- 6 domains
- 4 languages in development
- Size of the 6 domains:

	abdominal	poitrine	tête	accueil	suivi	habitudes
words	1'642	1'572	1'734	684	881	684
expansions for source	8'000'000	6'000'000	16'000'000	2'277	236'000	2'514
number of words in canonical patterns	620	551	604	110	217	66
canonical patterns	679	498	629	50	159	22

